Traceable and Comparable Evaluation Methodology for Biometric System Usability

by

Barbara Corsetti

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

Electrical, Electronics and Automation Engineering

Universidad Carlos III de Madrid

Advisor:

Raúl Sánchez Reíllo

Tutor:

Raúl Sánchez Reíllo

September, 2020

ACKNOWLEDGEMENTS

The work of this thesis was supported by a Marie Skłodowska-Curie EU grant within AMBER (enhAnced Mobile BiomEtRics) project. It was an honor and a privilege for me to have contributed to this project.

First of all, I would like to thank the person who made this possible: Thank you, Raul! Thank you, Raul, for giving me this opportunity, for teaching me, guiding me, and for all your valuable advice!

Then, since this Thesis let me live a lot of new adventures, I think I have many reasons to thank it…

Thanks to this Thesis, I spent three amazing months in Kent! Thank you, Richard, for hosting me in your group, for your support, availability, and your revisions. Thanks to my EDA colleagues who let me consider Canterbury one of my two homes far from my home. Thank you, Marco, for being there, for your support in writing even this Thesis!

Thanks to this Thesis, for the first time, I experienced working in a company. Thank you, Rames, for let me join the 11Paths group! Thank you, Unai, for your guide and help during this internship (and to update me on the pandemic in Italy)! Thanks to all my Telefónica colleagues for sharing with me ideas, launches, and shopping!

Thanks to this Thesis, I noticed how a simple PhD meeting can mean a lot. Thank you Dragos for adding me on LinkedIn! Thank you for the last two years and for the next ones (but please, give up eating Pesto!).

Thanks to this Thesis, I understood that, despite the distance, I always have the support of my family. Thank you, mamma!

PUBLISHED AND SUBMITTED CONTENT

Journal papers:

- Ramon Blanco-Gonzalo, Oscar Miguel-Hurtado, Chiara Lunerti, Richard M Guest, Barbara Corsetti, Elakkiya Ellavarason, Raul Sanchez-Reillo, "Biometric Systems Interaction Assessment: The State of the Art", IEEE Transactions on Human-Machine Systems 49(5), pp. 397-410, 2019
    - Published.
    - Role: collaborating in writing the paper.
    - Partly included in Chapter 3.
    - The material from this source included in this thesis is not singled out with typographic means and references
    - URL: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8743482

Conference papers:

- Barbara Corsetti, Ramon Blanco-Gonzalo, Raul Sanchez-Reillo, "User Interaction in Mobile Biometrics", 26th European Signal Processing Conference (EUSIPCO), pp. 543-547, 2018.
    - Published.
    - Role: designing the user interaction evaluation and writing the paper.
    - Partly included in Chapter 3.
    - The material from this source included in this thesis is not singled out with typographic means and references
    - URL: https://ieeexplore.ieee.org/abstract/document/8553284

- Barbara Corsetti, Raul Sanchez-Reillo, Richard M Guest, Marco Santopietro, "Face Image Analysis in Mobile Biometric Accessibility Evaluations", 2019 IEEE International Carnahan Conference on Security Technology (ICCST), pp. 1-5, 2019.
    - Published.
    - Role: designing the evaluation, performing the experiment and writing the paper.
    - Partly included in Chapter 6.

  - The material from this source included in this thesis is not singled out with typographic means and references
  - URL: https://ieeexplore.ieee.org/abstract/document/8888437

- Barbara Corsetti, Raul Sanchez-Reillo, Richard M Guest, "Ergonomics in Mobile Fingerprint Recognition Systems: A User Interaction Evaluation", 26th European Signal Processing Conference (EUSIPCO), pp. 543-547, 2018.
  - Published.
  - Role: designing the user interaction evaluation, performing the experiment and writing the paper.
  - Partly included in Chapter 5 and 7.
  - The material from this source included in this thesis is not singled out with typographic means and references
  - URL: https://link.springer.com/chapter/10.1007/978-3-030-51369-6_51

OTHER RESEARCH MERITS

Journal papers:

- Matthew Boakes, Richard Guest, Farzin Deravi, Barbara Corsetti, "Exploring Mobile Biometric Performance through Identification of Core Factors and Relationships", IEEE Transactions on Biometrics, Behaviour, and Identity Science 1 (4), pp. 278-291, 2019.

- Elakkiya Ellavarason, Richard Guest, Farzin Deravi, Raul Sanchez-Reillo, Barbara Corsetti, "Touch-dynamics based Behavioural Biometrics on Mobile Devices? A review from a Usability and Performance Perspective", ACM Computing Surveys (CSUR), 2020.

Conference papers:

- R. B. Gonzalo, B. Corsetti et al. Goicoechea-Telleria, "Attacking a smartphone biometric fingerprint system: a novice's approach," *2018* IEEE International Carnahan Conference on Security Technology (ICCST), vol. 675087, no. October, pp. 1–5, 2018.

Abstract

Biometrics is currently replacing traditional authentication mechanisms in more and more contexts. Biometric recognition allows us to identify ourselves quickly and easily without requiring any effort as when remembering secret codes or passwords. For this reason, more and more users choose to unlock their smartphones through the face or finger recognition. Hence, the release of mobile services supported by biometric recognition is increasing.

This growing use of biometrics has raised new questions about the user interaction. Is it easy to interact with the new biometric applications? Can anyone use them? Are people aware of how this new technology works? Can mobile biometrics be applied to improve accessibility in daily contexts?

This thesis aims to bring improvements to the study of user interaction and usability in biometrics. The study focuses on the user and his or her characteristics (e.g., experience, age, health status). For this reason, two scenario evaluations were carried out proposing biometric systems to a heterogeneous group of data subjects. The experiments were developed and designed following the directives of the ISO/IEC 21472. While the evaluation of the user interaction was conducted according to a novel methodology presented for the first time in this Thesis.

The novel methodology sets specific metrics to report how the user's characteristics impact the outcome of the biometric process. Its key point is the accessibility and all its aspects: the accessibility of the scenario, of the biometric system, and how the user's accessibility concerns influence the interaction with biometric applications.

Results show a strong correlation between the age, experience, health status of the user, and the outcome of the recognition process (in terms of performance, usability, accessibility, and biometric sample quality). This demonstrated that the accessibility brings to a lot of information regarding the human-biometric system interaction. Thus, all its aspects must be considered to improve the traceability and comparability of future assessments.

The main contributions of this Thesis include:

- First application of the ISO/IEC 21472 to design user interaction evaluations in Biometrics
- A proposal of a novel methodology to analyse the accessibility of biometric recognition systems
- Validation of this novel methodology through two user interaction evaluations.

Resumen

Actualmente la biometría está reemplazando los mecanismos de identificación tradicionales en más y más contextos. El reconocimiento biométrico nos permite identificarnos de forma rápida y sencilla sin requerir ningún esfuerzo, por ejemplo: recordar códigos secretos o contraseñas. Por esta razón, cada vez más usuarios optan por desbloquear sus teléfonos móviles mediante el reconocimiento facial o de huella. Por lo tanto, está proliferando la creación de servicios móviles basados en el reconocimiento biométrico.

Este uso creciente de la biometría ha planteado nuevas preguntas sobre la interacción del usuario con el sistema. ¿Cuánto de fácil es interactuar con las nuevas aplicaciones biométricas? ¿Son accesibles a todo el mundo? ¿Conoce la gente cómo funciona esta nueva tecnología? ¿Se puede aplicar la biometría móvil para mejorar la accesibilidad en contextos diarios?

Esta tesis tiene como objetivo aportar mejoras al estudio de la interacción del usuario y la usabilidad en biometría. El enfoque principal de este estudio es el usuario y sus características (por ejemplo, experiencia, edad, estado de salud). Por este motivo, se realizaron dos evaluaciones de escenarios proponiendo sistemas biométricos a un grupo heterogéneo de sujetos. Los experimentos se desarrollaron y diseñaron siguiendo las directivas de la ISO/IEC 21472. Mientras que la evaluación de la interacción del usuario

se realizó de acuerdo con una metodología novedosa presentada por primera vez en esta Tesis.

La nueva metodología establece métricas específicas para informar cómo las características del usuario impactan el resultado del proceso biométrico. Su punto clave es la accesibilidad y todos sus aspectos: la accesibilidad del escenario, del sistema biométrico y cómo los problemas de accesibilidad o discapacidad del usuario influyen en la interacción con las aplicaciones biométricas.

Los resultados muestran una fuerte correlación entre la edad, la experiencia, el estado de salud del usuario y el resultado del proceso de reconocimiento (en términos de rendimiento, usabilidad, accesibilidad y calidad de la muestra biométrica). Esto demuestra que la accesibilidad aporta mucha información sobre la interacción humano-sistema biométrico. Por tanto, todos sus aspectos deben ser considerados para mejorar la trazabilidad y comparabilidad de evaluaciones futuras.

Las principales contribuciones de esta Tesis incluyen:

• Primera aplicación de la norma ISO/IEC 21472 para diseñar evaluaciones de interacción del usuario en biometría

• Una propuesta de una metodología novedosa para analizar la accesibilidad de los sistemas de reconocimiento biométrico

• Validación de esta novedosa metodología a través de dos evaluaciones de interacción de usuarios.

List of Contents

List of Figures

List of Tables

List of Acronyms

ABC	Automated Border Control
ANOVA	Analysis of Variance
DET	Detection Error Trade-of
DHS	Department of Homeland Security
EER	Equal Error Rate
FAR	False Acceptation Rate
FMR	False Match Rate
FNMR	False non-Match Rate
FRR	False Rejection Rate
FTA	Failure to acquire
FTE	Failure to Enrol
HBSI	Human Biometric System Interaction
IEC	International Electrotechnical Commission
IP	Internet Protocol
ISO	International Organization for Standardization
NIST	National Institute of Standard and Technology
PCA	Principal Component Analysis
PoS	Point of Sale
SIFT	Scale Invariant Feature Transform
US – VISIT	United States Visitor and Immigrant Status Indicator Technology

Chapter 1 Introduction

1.1 Background

Nowadays, biometrics is a tool used in public and private contexts, where the identity of the users must be verified. Millions and millions of citizens are experiencing biometric recognition processes when they apply for passports or in border controls while traveling. At the same time, thanks to the latest mobile phones, biometrics is widely applied even in private contexts. For example, it is very common to unlock the screen of a smartphone using fingerprint or face identification. Besides, biometrics also supports the security of different mobile applications that contain private information of the users. Banking,

email, and payment apps ensure the privacy of the users precisely because their access is tied to a successful biometric authentication attempt.

It is impossible to establish the exact number of people who are currently using biometrics, but considering that in 2024 biometric recognition it will be used on 800 million mobile devices [1], we can estimate than in 4 years a large part of the world group will be using biometric recognition almost every day. Thus, we can foresee a growing trend of mobile biometrics in more and more daily contexts such as: accessing public services, paying at retail cash, or even opening home's doors. In these scenarios, mobile biometric authentication processes represent also an improvement of the usability and accessibility compared to many common gestures. Providing our credentials through biometric-based apps, instead of presenting badges or paying with cash and credit cards, could be more comfortable and faster, and it could break down many accessibility barriers.

Considering all these benefits in terms of usability, accessibility, and security, mobile biometrics should be available especially for all those people that have many difficulties when performing the above-mentioned tasks autonomously (for example, people with mobility and cognitive issues). To make this happen, firstly, it is necessary to establish which ones are the basic parameters for developing biometric applications that are universally accessible and user-friendly.

Depending on the characteristics of the user, the interaction with the biometric system and its performance could be affected by several external factors. Unlocking the phone with the fingers on a button sensor is a simple gesture, but it could be unwieldy to many groups of users that are affected by motor problems. At the same time, face recognition processes could be complicated to manage for people that are not used to interact with the smartphone's camera for taking selfie photos (e.g., elderly users and people with cognitive issues). These two examples represent two scenarios in which the capability and the experience of the users could compromise the use and the performance of biometric systems. For these reasons, it is important to conduct biometric system interaction assessments recruiting participants from different sectors: young people, older adults, mobility, and cognitive impaired users.

Understanding which accessibility barriers are the ones that prevent users from completing a successful biometric recognition process is the basis for developing the new generation of biometric applications; a new generation that must aim to be increasingly accessible and usable for all categories of users.

1.2 Objectives and Method of the Thesis

When people approach a biometric sensor to complete an identification process, there could be different factors influencing the performance of the recognition procedure. The evaluation of biometric systems is a complex topic especially because of all the external causes affecting the outcome of the authentication process. During the last decades, the user interaction was studied to understand the origins of these influencing factors and to evaluate the extent to which they affect the user experience and the matching scores [2], [3].

In the literature, the user interaction was mainly observed testing the usability, the performance, and few times reporting even the accessibility of biometrics systems. Through these metrics, researchers tried to establish whether the biometric systems are accessible and how easily people complete authentication processes successfully.

The work of this Thesis starts reviewing the main studies previously carried out regarding this research field. Once discussed the protocols and methodologies of these works, it will clear that, up to now, no specific methodology or metrics was proposed to evaluate the accessibility while testing the usability and the user interaction in biometrics.

In this Thesis, it will be proposed a formal methodology to evaluate the accessibility of biometric systems and how user's accessibility concerns influence usability and the outcome of the recognition process.

This novel methodology will be validated by examining the data collected during two experiments (fully included in this Thesis). Both experiments were planned and carried out following the same procedure articulated in 3 steps (Figure 1):

Figure 1: Three steps-procedure for performing the experiments.

- Design and development: in this part of the experiment, a context of use is recreated (e.g., an access control scenario or a retail place scenario). Therefore, a biometric recognition system is developed to be used in that context. The design of the scenario evaluations and the development of the biometric systems are completed following the guidelines of ISO/IEC 19795-2 [4] and ISO/IEC 21472 [5]. Since the smartphone environment is currently the main application field, the scenario evaluations reported in this Thesis are planned to evaluate the accessibility using mobile biometric solutions.
- Scenario Evaluation: throughout this part, the system is assessed by recruiting a group of volunteers with different characteristics and belonging to a specific sector of users (e.g., younger users, people with accessibility concerns, elderly people). Participants are asked to interact with the system completing biometric recognition processes.
- Analysis of the results is the last step of the experiment. At this point, the data collected during the scenario evaluation is analysed according to the formal methodology proposed in this Thesis.

1.3 Structure of the Thesis

To describe the whole work performed in this Thesis, this document is articulated in the following 8 chapters:

Chapter 2 defines the meaning of Biometrics, the main biometric modalities, and presents a summary of the History of biometric applications. Additionally, the chapter explains what a biometric recognition system is and details the ISO (International Organization for Standardization) [6] directives regarding the evaluation of biometric systems according to the performance and the sample quality.

Chapter 3 relies on the State of the Art in biometric usability and user interaction evaluations. The chapter starts with the definition of usability according to the international standards and follows presenting the first usability tests conducted by NIST (National Institute of Standards and Technology) [7] and by other research teams. Then, the user interaction evaluation is discussed describing the HBSI (Human Biometric System Interaction) model [8]. The final part of this chapter underlines the role of accessibility in user interaction evaluations. The importance of this aspect is explained by reviewing the previous works done to estimate the accessibility of biometric systems. Additionally, such estimation is included in current user interaction methodologies.

Chapter 4 presents a formal Methodology that aims to better report the accessibility while conducting user-biometric system interaction evaluations. This novel Methodology defines specific metrics to evaluate such as: a) the accessibility of the scenario (in which the user is required to complete a biometric recognition); b) the accessibility of the biometric system and, finally, c) the extent to which users accessibility problem affects the outcome of the biometric procedure (in terms of performance, usability, and sample quality).

Chapter 5 describes the Set-up of the Evaluations carried out to validate the formal Methodology proposed in Chapter 4. For both experiments, the system, the biometric devices, the workflow, and the data crew will be detailed after having clarified the ethical implication of the data collections.

Chapter 6 focuses on the evaluation of the first system: Access Control System through Biometric Recognition. The system was developed to help people to complete one of the most frequent daily actions: opening a door. The system is assessed through various scenarios in which the user performs fingerprint of face recognitions employing different biometric sensors (Figure 2).

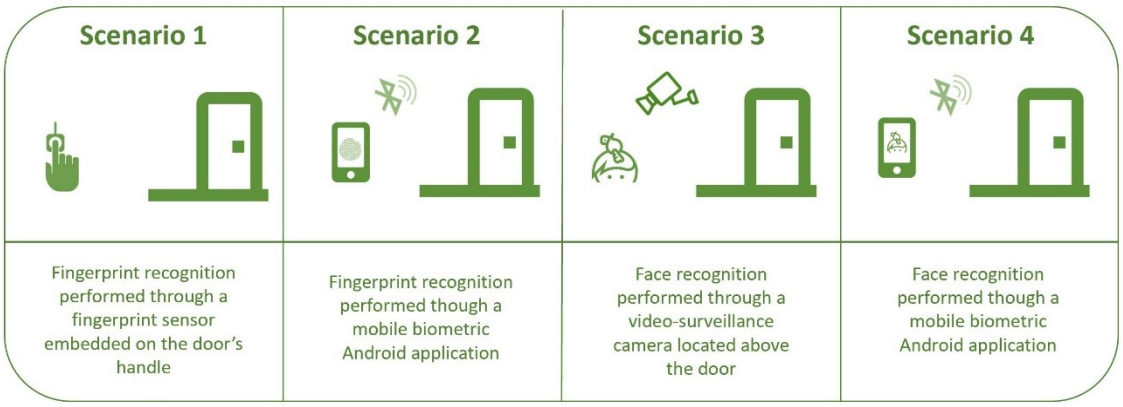


Figure 2: Access Control System through Biometric Recognition.

The data collected during the experiment will be reported following the novel methodology and the results will be gathered according to the user's group and the phase of the experiment.

Chapter 7 deals with the assessment of the second system: Mobile Fingerprint System for Retail Payments. This system is an Android mobile application based on fingerprint authentication: users can complete retail payments once the smartphone identified their fingerprint traits. The app was evaluated by the participants in three scenarios changing the design and the ergonomics of the smartphone's fingerprint sensor (Figure 3).

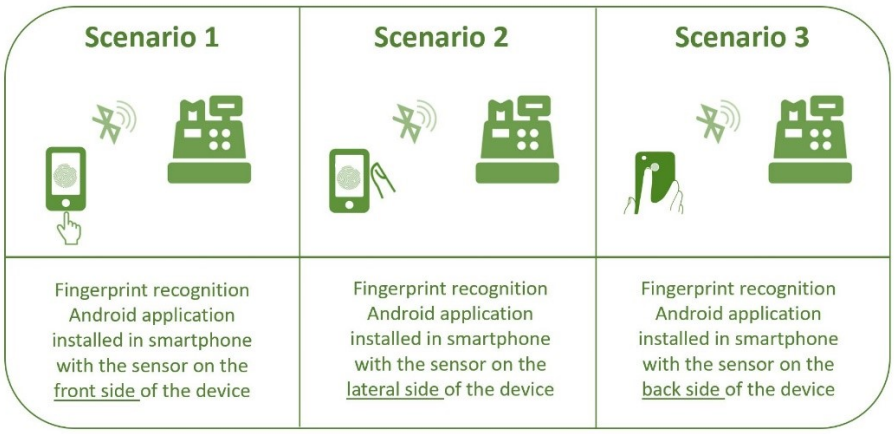


Figure 3: Mobile Fingerprint System for Retail Payments.

As the previous experiment, the data collected during the experiment will be reported following the novel methodology and the results will be gathered according to the user's sector and the phase of the evaluation.

Chapter 8 finally discusses the main contribution of this work reporting the Conclusions and Recommendations for Future Works.

Chapter 2 Biometrics

2.1 Definition of Biometrics and Biometric Modalities

In book XIX of the Odyssey, Homer described how Ulysses, after 20 years, was recognized by his old nurse Eurycleia. While she was washing him, the nurse noticed the scar Ulysses had on his leg since he was a child.

In the second part of the Divine Comedy, another masterpiece of the Literature, Dante hearing a soul singing in Purgatory recognized his friend Catella who died several years before his imaginary journey into the afterlife.

Our body and our behaviour define who we are, and they represent a tool through which other people recognize us. These two literary episodes are a good explication of

this concept and define what biometrics is. Biometric recognition is a process through which a human being is identified, analysing his physiological traits (like a fingerprint, a face, or even a scar), his behavioural characteristics (e.g., signature, keystroke, swipe), or other features that are a combination of physiological and behavioural traits (such as the voice).

Nowadays, many biometric features are applied in the identification processes, and they are mainly split into two main groups: physical and behavioural modalities as shown below (Figure 4).

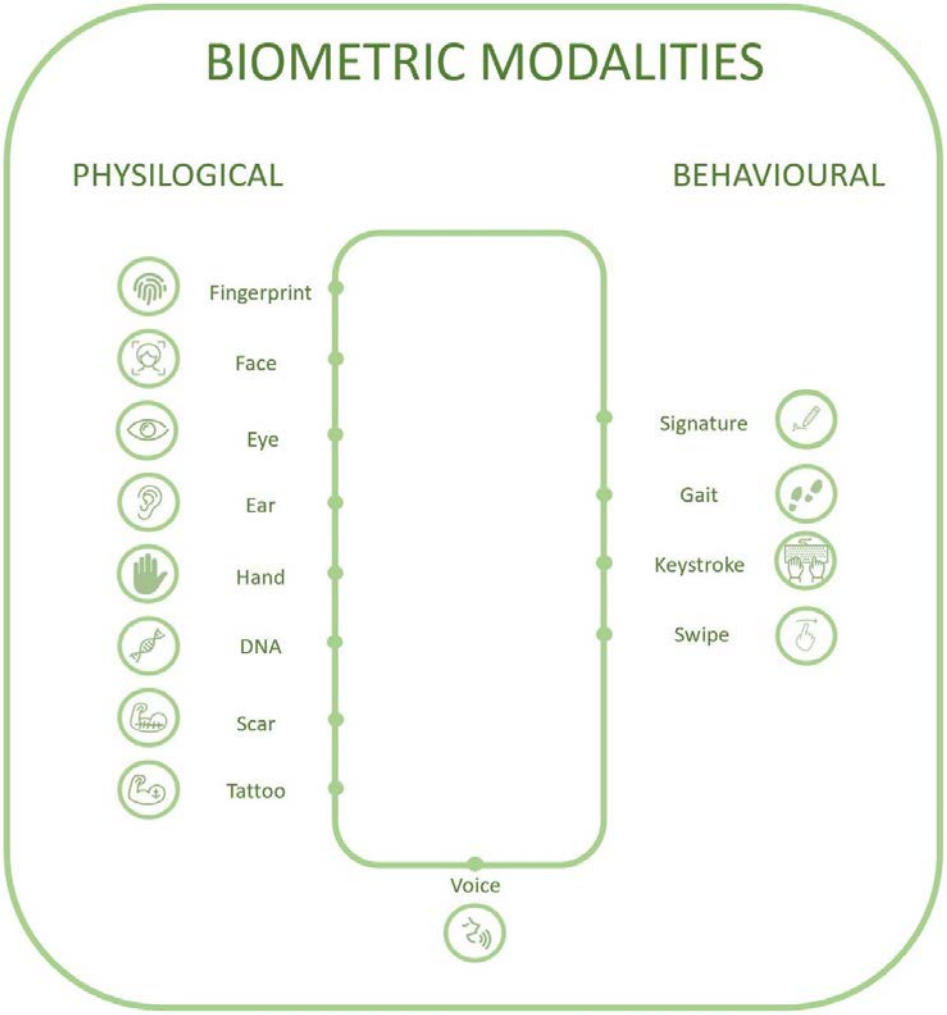


Figure 4: Biometric Identification Modalities.

All typologies of biometric features are widely applied in recognition mechanisms because they are universal (every human being owns biometric traits), unique (each biometric trait changes from person to person), collectible (biometric trait can be observed, stored, processed, and analysed).

The universality, uniqueness, and collectability are just some of the characteristics that made biometric traits the most sophisticated tools in the identification processes.

The experiments, that are going to be described among the next Chapters, evaluate two biometric systems based on the most applied biometric modalities: fingerprint and face recognition.

The fingerprint authentication is the biometric process that identifies a human being analysing the dermal features of his finger. The fingerprint contains a set of lines that can take on different configurations better known as minutiae (Figure 5).

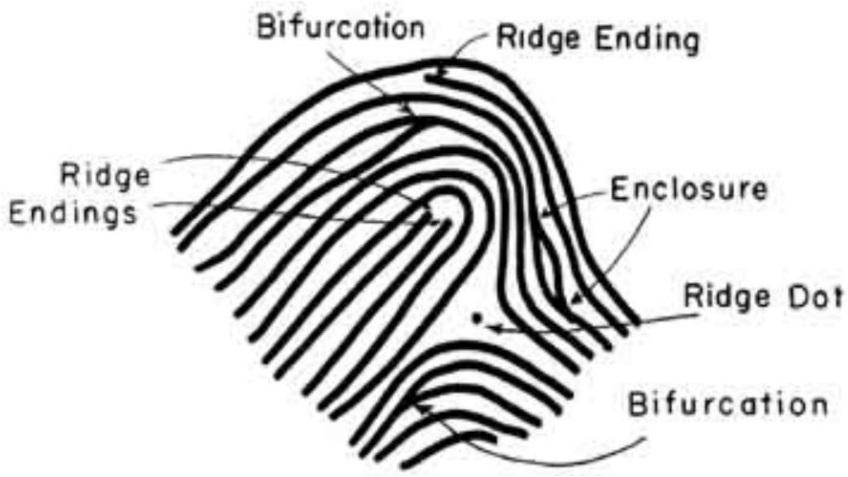


Figure 5: Fingerprint with the minutiae diagram [9]**.**

According to the current state of the art [10], the minutiae detection and comparison are the basing steps behind of all the fingerprint recognition algorithms.

The facial recognition, on the other hand, is based on analysing the principal features of a face (the shape and the size of the facial elements like the eye, the nose, the mouth,

and the geometry of the face). Due to the growing application of face recognition in daily scenarios, up to date, different approaches are proposed to develop high-accuracy facial recognition algorithms [11]. Generally, these algorithms can be gathered into 3 groups: local, holistic, and hybrid approaches.

The local approach algorithms consider just specific face features without considering the entire face region. An example of this approach is the SIFT (Scale Invariant Feature Transform) algorithm [12]. While the holist algorithms are based on an approach that takes into consideration all the facial region features (e.g., the Principal Component Analysis - PCA algorithm [13]). Finally, the hybrid approaches enclose all those algorithms that consider local and global characteristics of the face.

Through the next sessions, a snapshot of the main historical moments of Biometrics will be provided. Then, basic information about biometric systems and their evaluations will be discussed.

2.2 Historical Review of Biometric Applications: from 500 B.C. to the Face-ID

Although in lots of prehistoric caves there were found hundreds of paintings of hands and footprints (probably used to identify the painting's author), the first certain application of biometrics is due to the Babylonians around 500 BC. Thanks to this population, famous for the construction of the Ziggurats, there were the first collections of fingerprints in history. In fact, in many Babylonian tablets, used for commercial agreements, there were found fingerprints engraved as if they were signatures.

Fingerprint traits were widely applied as a signature in business contexts. Many historical proofs certify its use also among various populations coming from China and Egypt. But, even if fingerprints were used since almost the beginning of human history, they started to be seriously studied just at the end of the 1600s thanks to Nehemian Grew, Govad Bidloo, and Marcello Malpighi. These three scientists intensively observed the anatomy of the human body and reported, for the first time, the anatomical characteristics that are present on the fingers surface (e.g., spirals and ridges) (Figure 6).

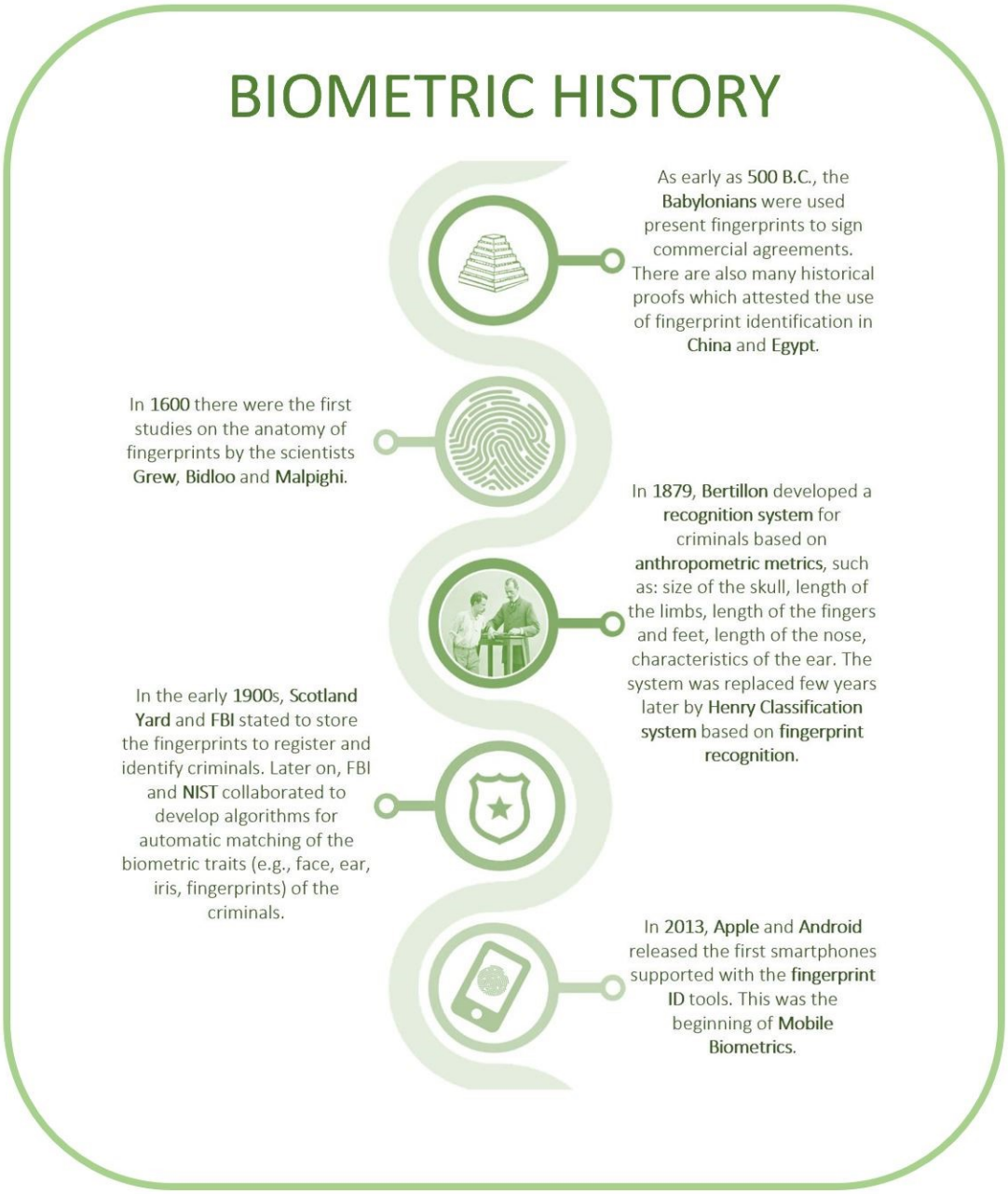


Figure 6: Main moments of Biometrics history.

Later, Alphonse Bertillon, a French criminologist, proposed a system for collecting the physiological characteristics of the criminals. The Bertillon system, dated 1879, was based on 5 anthropometric measurements: the size of the skull, length of the limbs, length of the fingers and feet, length of the nose, and the characteristics of the ear. This system

was used in forensics for a few decades between the late 1800s and the early 1900s when it was replaced with recognition systems based on fingerprints.

The application of fingerprints for forensic purposes is due to Sir Francis Galton, a British explorer and anthropologist, who for long studied the probability that two or more people could have the same fingerprint. He introduced, for the first time, the term of minutia and demonstrated the uniqueness of the fingerprint traits. Later, he started to collaborate with the British policeman Sir Edward Henry. Analysing a database of criminals' fingerprint samples collected by Henry in India, they set the criteria for identifying those samples that belonged to the same human being. This method for the classification of fingerprint characteristics is still known as the Henry Classification System. Based on this classification system, in 1900, Scotland Yard (British police) proposed its apparatus for the identification of criminals employing fingerprints.

In 1903, the FBI, the Federal Investigation Office of the United States of America, also began to register criminals through their fingerprint traits. From that moment, more and more interest arise in the process of identifying people using biometric features. Besides the use of fingerprints, during the last century, there were developed the first systems based on the recognition of the voice, iris, and signature for biometric identification applications.

At the same time, the use of fingerprint recognition apparatus by the FBI became increasingly sophisticated and, to obtain better results, FBI established the collaboration with the NIST (National Institute of Standard and Technology). These two institutions worked together to develop an automated system for capturing and matching of the criminals and their fingerprints. This collaboration produced the M40 algorithm, the first matching algorithm used by the FBI to identify criminals. After that, more sophisticated matching algorithms were developed not only for fingerprints but also for facial, voice, iris, and signature traits. Besides forensics, biometrics started to be applied in other public contexts such as border check points, access controls, banking, or when requesting a passport or an ID card.

In 2002, another milestone in the history of Biometrics was reached: the ISO (International Organization for the Standardization) created the first subcommittee for

standardizing biometric systems (ISO / IEC JTC1 SC37 [14]). Even today, the SC37 aims to define useful guidelines to evaluate a biometric system under different aspects (e.g., quality of the sample, usability, performance, and security).

Later on, in 2013 Apple introduced fingerprint recognition as an unlocking mechanism for the iPhone 5S [15]. This moment can be defined as the birth of the Mobile Biometrics. From that moment lots of biometric applications were developed to improve the security in more and more private scenarios. The use of biometrics as mobile privacy settings was then intensified with the release of the Face ID on iPhone devices in 2017 [16].

2.3 Biometric Systems: Functioning and Evaluations

A biometric system is a device adapted to perform an authentication process which is generally composed of two distinct phases: enrolment and verification.

The enrolment starts when a user presents his physical or behavioural characteristics to a biometric sensor that can capture and store them. An example of this phase is when, for the first time, we register fingerprints on our smartphone. In this case, the security settings of the device require us to interact with the fingerprint sensor several times by presenting different finger portions (e.g. central, lateral, upper, and lower side). This procedure is necessary because while we are interacting with the sensor, the system extracts as many features as possible from our fingerprint and creates a template that is encrypted and stored in the background of the smartphone.

The verification part, on the other hand, consists of a matching process that is implemented between a biometric template and another biometric sample. Returning to the case of the smartphone, the verification phase occurs all those times that we attempt to unlock the mobile phone by touching the biometric sensor with our finger. When we touch the biometric sensor, the system compares the finger image provided on the spot with the stored template obtaining a matching score. If this matching score is higher than a threshold value, the authentication is established (genuine attempt) and the mobile

phone is unlocked. Otherwise, if the matching score is lower than the threshold value the identification is rejected (impostor attempt) and the process should be repeated.

Biometric systems can be evaluated under different aspects such as performance, usability, vulnerability, security, and quality of biometric data. This Thesis deals with the evaluation of all those factors that, depending on the user's capabilities, modify the outcome of the biometric process. Thus, the focus of the work will report information regarding performance, usability, and even biometric data quality.

The evaluation of the usability in biometrics will be discussed in detail alongside the State of Art's Chapter (Chapter 3) of this Thesis. While useful notions regarding the methodologies to assess the performance and the quality of the biometric data will be provided in the next subsections.

2.3.1 Biometric Systems Performance Evaluation

The International Organization for Standardization sets guidelines to evaluate the performance through the multi-part ISO/IEC 19795 [17] standards (Table 1).

Table 1: Summary of the International Organization for Standardization sets guidelines to evaluate the performance of biometric systems.

Standard Identifier	Title
ISO/IEC 19795 - 1 [18]	Principles and framework
ISO/IEC 19795 - 2 [4]	Testing methodology for technology and scenario evaluation
ISO/IEC 19795 - 3 [19]	Modality-specific testing
ISO/IEC 19795 - 4 [20]	Interoperability performance testing
ISO/IEC 19795 - 5 [21]	Access control scenario and grading scheme
ISO/IEC 19795 - 6 [22]	Testing methodologies for operational evaluation
ISO/IEC 19795 - 7 [23]	Testing of on-card biometric comparison algorithms
ISO/IEC 19795 - 9 [24]	Testing on mobile devices

In Part 1, Principles and Framework [25], there are described the 3 kinds of evaluation through which it is possible to assess the performance of a biometric system. This means technology, scenario, and operational evaluation.

The technology evaluation is the process to assess the performance of specific recognition algorithms using a pre-collected biometric samples database or conducting a new data collection [26], [27], [28].

The Scenario Evaluation refers to the procedure in which the prototype of a specific biometric system is developed and tested by data subjects in a simulated scenario [29], [30].

The operational evaluation assesses the performance of a biometric system in a real application scenario recruiting a real group of customers. An example of this evaluation is to assess the outcome of biometric recognition processes performed by travellers at Automated Border Control (ABC) gates [31].

The quantitative evaluation of the biometric system performance is carried out evaluating specific error rates as:

- FTE is the Failure-To-Enrol rate of people who provided their biometric traits during the enrolment to the sensor, but the system was not able to store a template.

- FTA is the Failure-To-Acquire rate of biometric samples that cannot be acquired by the sensor during the enrolment and verification.

- FMR is the False-Match-Rate of impostor verification attempts that are classified as genuine attempts

- FNMR is the False-Non-Matching rate of genuine verification attempts classified as impostor attempts.

In case we are analysing verification transitions (set of various attempts), we must report the following performance metrics:

- FRR is the False-Rejection-Rate that indicates the rate of genuine verification transactions that are classified as impostor transactions. The FFR could be calculated as:

$$FRR = FTA + FNMR\,(1 - FTA)$$

When the verification transactions comprise of just one verification attempt.

- FAR is the False-Acceptation-Rate that indicates the rate of impostor verification transactions that are classified as genuine transactions. The FAR could be calculated as:

$$FAR = FMR(1 - FTA)$$

When the verification transactions comprise of just one verification attempt.

- EER (Equal Error Rate) is the point where the proportion of FMR is equal to FNMR. The lower values of EER indicates a higher level of the system's performance.

Additionally, the standard ISO/IEC 19795 - 1 [25], suggests a modality to plot the performance of a biometric recognition system:

- DET (Detection Error Trade-off) is a curve that plots the FRR in the y-axis and the FAR on the x-axis. In this case. Thus, the curve represents the false-negative vs. the false-positive.

2.3.2 Biometric Sample Quality Evaluation

By analysing the quality of biometric data collected during a system evaluation, we can extract a lot of information. Sample quality could depend on the accuracy of the biometric sensor, or the users, or even on the environmental conditions. Since low-quality samples impact the performance of the comparison process is recommendable conducting quality assessments while testing a specific biometric system.

Standardized metrics to evaluate quality data are provided through the ISO/IEC 29794 [17] whose parts 4 and 5 are specific for finger and face samples (Table 2).

Table 2: Finger and face sample metrics standardized in ISO/IEC 29794.

Standard Identifier	Metrics to report the quality of the biometric data
ISO/IEC TR 29794-4:2010 Information technology - Biometric sample quality - Part 4: Finger image data [32]	- Region of interest (ROI) - Minutiae extraction
ISO/IEC TR 29794-5:2010 Information technology - Biometric sample quality - Part 5: Face image data [33]	- Subject's behaviour (e.g., closed and open eyes, closed and open mouth, any kind of expression, head pose) - Brightness and background analysis

Regarding the fingerprint quality analysis, the National Institute of Standards and Technology (NIST) developed the NFIQ (NIST Fingerprint Image Quality) algorithm [34] based on the recommendation provided in [30]. This algorithm can extract the minutiae from a finger image sample returning the number of minutiae, the typology (e.g., ridge or bifurcation) assigning a quality value to each minutia. Besides, the algorithm associates a specific NFIQ value to each finger sample to indicate the overall image quality level. NFIQ = 1 is the highest level indicating a good-quality finger image, while NFIQ = 5 indicates the lowest quality samples.

During the last decade, the NFIQ score is widely applied in fingerprint quality evaluations [35].

Recently, NIST update this algorithm launching and new version: NFIQ2. This version guarantees more accuracy in the analysis of the fingerprint sample quality although it is trained just to assess images captured by optic sensors or scanned from inked cards [36].

Chapter 3 The State of the Art

This chapter starts reviewing the main usability and user interaction evaluations carried out in Biometrics. By discussing the methods, the methodologies, and the findings of these works, the aim is to establish the improvements to bring in this research area.

For this reason, the two last subsections provide a discussion of the accessibility evaluations done in biometrics explaining why the accessibility should be included in Biometrics while testing the usability and user interaction.

3.1 Usability Evaluation in Biometrics

When biometric recognition processes began to be applied in more and more contexts (e.g., border checks, banking, access controls), arose the need to evaluate how people interacted with biometrics in those circumstances. Thus, several researching groups started to analyse the usability metrics while evaluating biometric systems. Since the ISO 9241-11:1998 [37] defines usability as "*the extent to which a product can be used by*

specific users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use", in Biometrics the usability evaluation means assessing how easily users complete a successful identification process reporting:

- The effectiveness that generally is indicated as the number of incorrect interactions or errors made by the users during a biometric recognition procedure.
- The efficiency that is a temporal metric indicating the time spent by the user to complete a recognition process.
- The satisfaction which is the metrics that assess the users' opinion regarding the interaction with a biometric recognition device.

Hence, efficiency and effectiveness are quantitative metrics instead of the satisfaction that represents a qualitative measure of usability.

Less often, researchers evaluate the usability including also the:

- Learnability that investigates the users' ability to understand how a specific recognition system works.
- Memorability that studies the users' ability to remember how a specific recognition system works.

3.1.1 Usability Tests by NIST

During the last 15 years, the National Institute of Standards and Technology (NIST) strongly promoted the evaluation of the usability assessing biometric recognition systems in the emerging application scenarios. In 2005, NIST founded the Visualization and Usability Group [38] that was the first group to provide a taxonomy of definitions to be used in biometric usability evaluations [39]. At the same time, this group started carrying out several experiments to assess how different factors (e.g., the gender, the age, or the users' habits [40]) could influence the usability of biometric systems. However, the main contribution of this research team was the publication of the handbook "Usability and Biometrics: Ensuring Successful Biometric Systems" [41] that describes the NIST Usability Model (Figure 7).

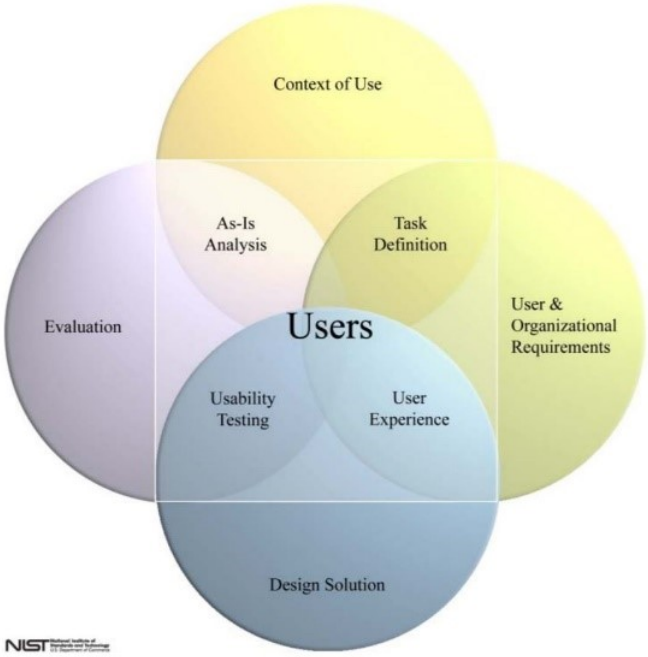


Figure 7: NIST Usability Model [41]**.**

This model is designed to develop easy-to-use and high-performance biometric systems establishing a design procedure based on the following 4 steps:

- Defining the Context of Use: in this part, the focus is the setting and preparation of the scenario in which the system is supposed to be used. Moreover, special attention is given to the task that the user must complete interacting with its biometric sensor.
- Determining the User and Organizational Requirements, which means establishing user requirements, environmental requirements, and technical requirements.
- Developing the Design Solution: this part deals with the establishment of the system layout, the user interface, and the materials.
- Conducting the Evaluation: during this phase, the user interaction with the biometric system is assessed in terms of usability, accessibility, and performance.

Over this model, the work of NIST regarding the usability evaluation in biometrics enclosed several studies (Table 3), carried out by collaborating with the American Department of Homeland Security (DHS) [42] and the United States Visitor and Immigrant Status Indicator Technology (US-VISIT) [43] program.

Table 3: Main findings in NIST usability evaluations.

Publication	Evaluation Type	Biometric System	The protocol of the Experiment and Analysed Metrics	Findings
Effects of Scanner Height on Fingerprint Capture [44]	Technological Evaluation	(US-VISIT) Fingerprint Scanner	Usability evaluation of a fingerprint scanner collocated at different heights. Users received verbal instruction before the evaluation. Usability was reported in terms of efficiency, effectiveness, and satisfaction).	Better results in terms of efficiency and effectiveness were obtained at specific heights. There is a clear dependence between the user satisfaction and the scanner height.
Usability Testing of Ten-Print Fingerprint Capture [45]	Technological Evaluation	US-VISIT Fingerprint Scanner	Users interacted with a ten-print finger capture device, receiving instructions through different means: poster, video, and verbal. Results were reported in terms of efficiency, effectiveness, and satisfaction.	Receiving verbal and video instruction helps the users experience while reading the instruction on a poster made the user spend more time to complete the task. The help of the operators was necessary to complete the task successfully.
Usability Testing of Height and Angles of Ten-Print Fingerprint Capture [46]	Technological Evaluation	US-VISIT Fingerprint Angled Scanner	The system was tilted several degrees and several heights were used. Results were reported in terms of usability metrics: efficiency, effectiveness, and satisfaction.	Changing the height and the angles did not affect significantly the efficiency and effectiveness. There is a dependency between volunteer height and user opinions. Shorter participants declared the flatter angle (10º) as the less comfortable, while taller participants indicated the 30º as the less comfortable.
Usability Testing of Face Image Capture for US Ports of Entry [31]	Operational Evaluation	US-VISIT MBARK software for the storage of face samples	Users were asked to provide facial photos once they arrived at a control desk. The sample images were captured by an operator through an SLR camera. The operator received the preview of the image on a screen to evaluate the quality. The evaluation of the samples' quality was conducted by using a face overlay.	The use of an overlay in the evaluation samples' quality helped to obtain face image more centred and with a more appropriate pose of the user.
Assessing Face Overlay [47]	Technological Evaluation	One commercial webcam used to capture face samples	Participants were asked to capture "the best passport picture in the shortest amount of time" of a mannequin. The usability metrics were evaluated by reporting efficiency effectiveness and satisfaction.	Authors argued that the use of a face overlay helps to improve the face image quality. In fact, in terms of effectiveness, 53,2% of the face samples were centred and the other 45,4% partially centred.
Usability testing of a contactless fingerprint device: Part 1 [26]	Technological Evaluation	One contactless fingerprint scanner and one contact fingerprint scanner	Participants were asked to interact with both devices in 3 different contexts: without receiving instruction, receiving verbal instruction, and receiving video instruction. Usability was reported through the efficiency the effectiveness and satisfaction.	Participants declared their preference for the contact devices, probably due to a higher level of intuitiveness respect to the contactless system. Thus, NIST highlighted the necessity of educating people to properly use contactless scanners.
Usability testing of a contactless fingerprint device: Part 2 [27]	Technological Evaluation	One contactless fingerprint scanner and one contact fingerprint scanner collocated in a surface at 20 degree	Participants were asked to interact with both devices in 3 different contexts: without receiving instruction, receiving verbal instruction, and receiving video instruction. Efficiency, effectiveness, and satisfaction were reported to evaluate the usability of the systems.	There were many interaction issues with the contactless device. Even when users were required to interact with the contactless, they touched the scanner surface. Considering the results, as the first part, NIST underlined the importance to instruct people on how to interact with the biometric systems.
Contactless fingerprint devices usability test [28]	Technological Evaluation	Three contactless fingerprint scanners and one contact fingerprint scanner	Usability evaluation of 3 contactless devices in 3 different contexts: without receiving instruction and receiving verbal or video instruction.	The contactless scanners are viable biometric systems to be applied in border check scenarios, but people need more instruction regarding their use.

Thanks to the findings obtained from these studies, NIST contributed to the definition of specific guidelines to improve the usability and performance of biometric identifications in border control scenarios.

These studies found out: how to better place the fingerprint sensors [44], the most effective way to instruct travellers in presenting their biometrics traits [45] and, even, the capture software' specifications to help the border operator in taking high-quality biometric samples [47].

NIST findings also underlined the importance to inform people regarding biometric technologies [28] to let people more confident while performing an authentication. Thus, NIST's works were important not just for the researching community but even for the entire society.

3.1.2 Testing the Usability in Private Scenarios

Besides the NIST evaluations (mostly focus on the border control scenarios), other researching teams assessed the usability of the biometric systems (Table 4) applied as support in private tasks (e.g., banking, e-payment, security settings) even considering the mobile environment.

The publications, listed in table 4, tried to establish more comfortable configurations to allow users to complete biometric recognitions evaluating the usability of different scenarios [48], [49].

At the same time, the authors investigated the perception of people regarding the application of biometrics in daily contexts [29].

This was important to understand people's preference regarding the modalities in which provide biometrics traits and the users' willingness to perform biometric recognition process in daily scenarios.

Table 4: Relevant findings of usability evaluations.

Publication	Evaluation Type	Biometric System	The protocol of the Experiment and Analysed Metrics	Findings
Usability Evaluation of Voiceprint Authentication in automated telephone banking: Sentences versus digits[29]	Scenario Evaluation	Voice authentication using digits or sentence for banking	Users enrolled in the experiment were required to interact with an automated phone banking service and to provide voice traits repeating digits and sentences. The usability was assessed through a satisfaction questionnaire.	Volunteers found voiceprint authentication based on digits more comfortable compared with the voiceprint authentication based on sentences.
Usability analysis of a handwritten signature recognition system applied to mobile scenarios [30]	Scenario Evaluation	Handwritten signature on mobile devices (2 tablets, 1 smartphone, and 1 digitizer).	The 20 volunteers were required to provide their handwritten signature interacting with the mobile devices in 5 different scenarios representing the most typical posture in which people are used to interacting with mobile devices. The evaluation was divided into 3 sessions 7 days apart.	The authors reported an improvement in usability and performance between sessions. This demonstrated that more experience and training help in completing successful recognition attempts and improve performance.
Usability evaluation of biometrics in mobile environments [48]	Scenario Evaluation	Handwritten signature on a tablet device using 3 styluses with different tips' shape, length, and diameter.	The evaluation was divided into 3 sections 7 days apart from each one. Users interacted with the tablet device and 3 styluses in 3 different scenarios: seating with the table on a table, seating while holding the device, standing while the table was put on a tilted surface. Results were reported in terms of usability and performance.	Efficiency, effectiveness, and learnability improved during the 3 sessions. Users got experience among the different parts of the experiment. While performance score changed depending on the postures and the styluses demonstrating that ergonomics influence on system's outcome.
Usability Analysis of a Novel Biometric Authentication Approach for Android-based Mobile Devices [49]	Scenario Evaluation	Fingerprint scanner running on an Android smartphone	100 users were required to interact with a fingerprint recognition scanner embedded on a smartphone. A successful recognition process allowed the user to access a specific mobile application. The data collected during the interaction user-biometric system was analysed in terms of usability and performance.	Analysing users' opinions authors argue that interaction time is an important factor to be analysed. A lot of users declared that the task required to evaluate the system was quite slow and boring.

3.2 Human Biometrics Systems Interaction (HBSI) model

The usability evaluation was strongly introduced as a fundamental metric in the Human Biometrics Systems Interaction (HBSI) model [8]. In fact, in the HBSI framework (Figure 8), the usability was combined with the ergonomics and the sample quality to evaluate the outcome of biometric authentication devices.

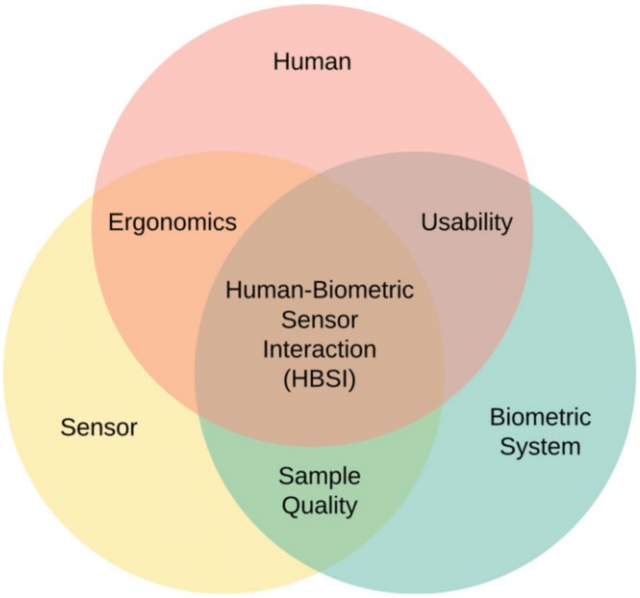


Figure 8: Human Biometrics Systems Interaction (HBSI) model [50]**.**

This model, proposed by Kukula and Eliot, has been validated by conducting several user interaction evaluations. In 2010 the HBSI framework was applied for assessing the user interaction with 3 swipe-based fingerprint sensors [51] and, later on, with a hand-geometry recognition system [52]. These studies underlined the importance of this model, especially, to gain better knowledge regarding the origins of the errors that users made when presenting their biometric traits to the recognition sensors.

The HBSI framework has been also applied in the smartphone scenario when, in 2015, authors evaluated a mobile voice recognition system [53]. 27 volunteers took part in this experiment testing the system during 2 sessions split in 2 different weeks. Following the HBSI guidelines, researchers assessed the usability metrics noticing that the users did less

incorrect interactions and completed the voice recognition more quickly in the second part of the experiment. This was probably due to the experience that the people gained from one section to another. The system proposed in this study was modified in a multimodal biometric application when, in 2016, authors included face recognition [54]. The voice and face system were evaluated through a 3-section evaluation. From the users' feedbacks received at the end of the whole experiment, face-recognition was rated more positively compared with the voice recognition. Besides, as the previous study, during the last sessions, users completed the tasks quickly and with fewer errors. These results showed that there is a dependence between user training and usability scores.

3.3 Accessibility: Improvement Point in User Interaction Evaluation

Although models and methodologies, described in the previous sessions, consider the user as one of the fundamental factors in the human-biometric system interaction assessments, few studies fully considered all the characteristics of the users recruited in the experiments. For instance, in Biometrics, just a limited number of studies tested the user interaction focusing on the accessibility level of the participants recruited in the experiments.

According to [55], the accessibility is "the *extent to which products, systems, services, environments, and facilities can be used by people from a population with the widest range of characteristics and capabilities, to achieve a specific goal in a specific context of use".*

As specified by the words: "*characteristics*" and "*capabilitie*s" a human being could be affected by different types of accessibility concerns (e.g., visual, mobility, auditory, and cognitive). For accessibility evaluations, carried out for the development of this Thesis, were enrolled users affected by mobility and cognitive issues.

Mobility disorder mainly affected the motor function of a human being. While the cognitive concerns cause different kinds of deficits to the intellectual abilities of a user. Cognitive issues are grouped in Developmental (congenital related disorder affecting the

cognitive abilities of a human being) and Learning Issues (attention deficits causing intellectual disorders).

In 2013 Sanchez-Reillo et al. [56] published a work on the accessibility in biometrics presenting one by one the main pathologies that can invalidate users, physically and cognitively. Thaking into account that these pathologies compromise the use of a specific biometric modality, authors proposed for each accessibility concerns the most suitable biometric recognition solution taking into consideration various contexts of use.

NIST conducted one of the first accessibility studies [57] when, in 2008, enrolled 12 users with visual concerns for assessing the usability of a fingerprint scanner. During the interaction, participates were guided receiving audio and vibration feedback while presenting the fingers to the sensors. Results have shown that the audio tones are the most efficient way to instruct blind people on how to interact with fingerprint recognition devices.

Later on, in [58] authors recruited 21 participants for evaluating the accessibility of a payment mobile application based on the signature and fingerprints recognition. The application allowed users to complete money transactions once the user has authenticated his/her fingerprint and handwritten signature. The 21 subjects, gathered as data crew, were affected by different physical pathologies (e.g., 10 volunteers had hands and arms concerns and 11 legs-motor problems). The assessment brought to the conclusion that through signature recognition users obtained better scores in terms of performance.

A part of the motor and the cognitive concerns, some accessibility evaluations were carried out also recruiting older users [59], [60]. This is since several accessibility pathologies are age-related and, additionally, aging affects the user's dexterity and cognitive skills which influence their ability to provide high-quality biometric traits [35]. For instance, as demonstrated in [60] and in [61] the application of fingerprint traits brought a low level of recognition performance compared to other biometric modalities.

In 2017, authors in [62], evaluating the accessibility of a mobile biometric system reported:

- *The number of test subjects who could not begin interaction with a modality*

- *The number of test subjects who could not complete the section*

These two scores represent the first metrics applied specifically to evaluate the accessibility of a biometric recognition system. During the evaluation, users were asked to interact with an Android app that allowed users to withdraw money from an ATM (Automatic Taller Machine) once authenticated using biometric recognition (face and fingerprint) and pin and pattern verification. The experiment was divided into two sessions 7 days apart. The 41 volunteers recruited in the experiment were split into two groups: accessibility (users with cognitive and motor concerns) and control (no accessibility problems) group. The authors noticed that all the users belonging to the control subgroup were willing to start the 5 modalities to complete the task required in the experiment. While for accessibility users were not possible to interact with some modalities (voice, face, and pattern) because of their accessibility concerns (especially visual and cognitive problems). Additionally, a significant number of users coming from both groups were not able to complete PIN and pattern in the second session because they forgot their credentials provided during the first visit.

Most of the works presented in this section were conducted in the same laboratory in which this thesis was also carried out. Thus, these previous evaluations were the basis on which this thesis was developed.

3.3 Conclusions

Through the discussion of the previous works, it is clear that the accessibility tests in Biometrics are currently carried out assessing the user interaction gathering elderly users or people with accessibility issues without reporting any specific metrics.

Even if biometric solutions are widely applied in more and more daily scenarios, there is no international standard providing metrics to evaluate the accessibility of a biometric system. While, in biometrics, accessibility evaluations are important to understand the reason why users have no access to a specific recognition system.

Thus, in the next chapter, a formal methodology to report the accessibility in biometrics will be described. Therefore, this methodology will be validated analysing the results of two experiments to be included in the evaluation of the user interaction.

Chapter 4 Novel Methodology

The following chapter provides the proposed formal methodology to evaluate the accessibility in biometric user interaction assessments.

The starting point for the development of this methodology was the ISO/IEC 21472 [5]. Even if this standard is not related to the accessibility (it is more focused on reporting the influence of the user interaction on the system performance), it provides useful guidelines to organize the set-up of the biometric evaluations. Therefore, the chapter starts by presenting the ISO/IEC 21472 and discusses its recommendations.

Later, the chapter describes a novel methodology to report the accessibility in biometrics. The methodology outlines three specific aspects that must be analysed in biometric accessibility evaluations. For each one of these aspects, the methodology provides specific metrics to assess each one of them.

4.1 The ISO/IEC 21472

The International Organization for Standardizations recently defined a new methodology to evaluate the user interaction influence on the biometric system's performance.

This methodology, reported in ISO/IEC 21472, provides the guidelines to design biometric evaluations and to report the effect of the user interaction on the performance of the recognition process.

This standard suggests designing more scenario evaluations to test how users interact with a biometric system. Specifically, the methodology states that the first scenario must be designed under Reference Evaluation Conditions (REC), meaning that it is the baseline scenario of the whole experiment. While other scenarios must follow the Target Evaluation Conditions (TEC) (Figure 9).

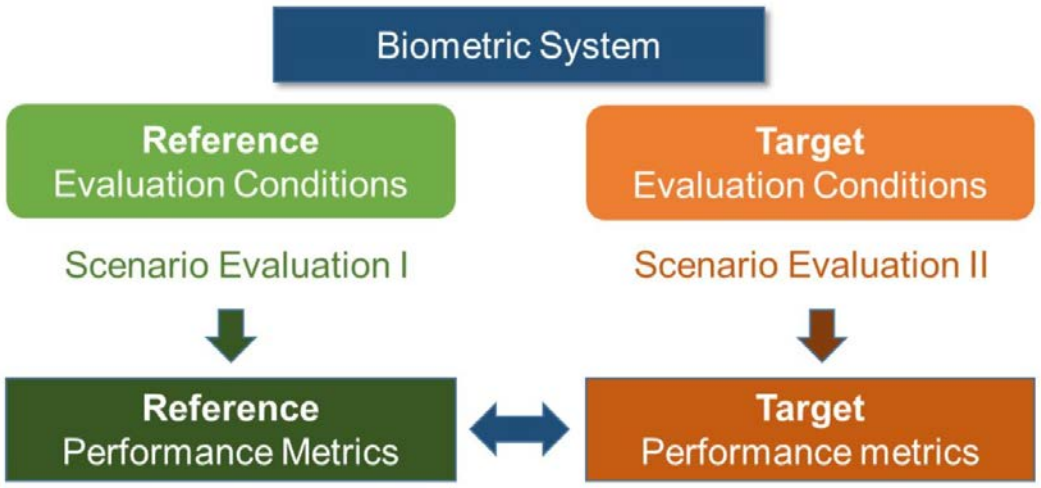


Figure 9: ISO/IEC 21472 [5].

It is recommendable to design the evaluation including a REC scenario and at least one TEC scenario. The difference between the baseline scenario and the TEC scenario relies on the user interaction factor to evaluate.

The user interaction factors are related to three aspects:

- the configuration, ergonomics, typology of the biometric system, and/or the biometric sensor
- the data subject characteristics (e.g., user skills or the biometric trait they must provide)
- the elements that influence the interaction between the biometric sensor and the users (e.g., guidance, training, or feedback).

According to this standard, the interaction with a biometric system should be evaluated through more scenarios employing different biometric sensors, recruiting more test subject groups, or changing the guidance tools.

The two experiments, reported in this Thesis, were planned by following these scenario evaluation settings. Hence, for both experiments, a specific biometric system was developed and tested in different scenarios. The recognition process was carried out through biometric sensors with various configurations and based on different biometric traits. The data crew of both experiments gathered different groups.

Besides reporting the effect of the user interaction on the system performance, these works aimed to report the information regarding the accessibility of the biometric systems. Since there are neither guidelines in the ISO/IEC 21472 nor other standards to evaluate the accessibility in biometrics, in the next sections a formal methodology containing accessibility metrics will be presented.

4.2 Accessibility Methodology

In the previous chapter, we read the ISO definition of accessibility [55] (Chapter 3.3). The standard defines accessibility through many terms such as the context of use, system, and the possibilities and capabilities of the users. These words are important because they explain the different meanings of the accessibility.

In fact, the accessibility may refer to the scenario, to the system, and the health status of the users. All three aspects must be considered in case we want to report the accessibility of a certain system. Regarding biometrics, the accessibility should be observed during user interaction assessments and it should be reported by:

- evaluating the extent to which users have access to the evaluation scenario.
- evaluating the extent to which a biometric system is accessible for the users and how easily they complete a recognition process.
- studying how the accessibility concerns of the users affect the whole outcome of the recognition process (in terms of performance, usability, and sample quality).

Based on these three points we are going to define a methodology to report the accessibility while testing the user-biometric system's interaction.

It is a formal methodology that aims to set guidelines and metrics for testing the accessibility of the scenario, of the system, and for reporting how motor or cognitive concerns of the user influence the outcome of the biometric procedure (Figure 10).

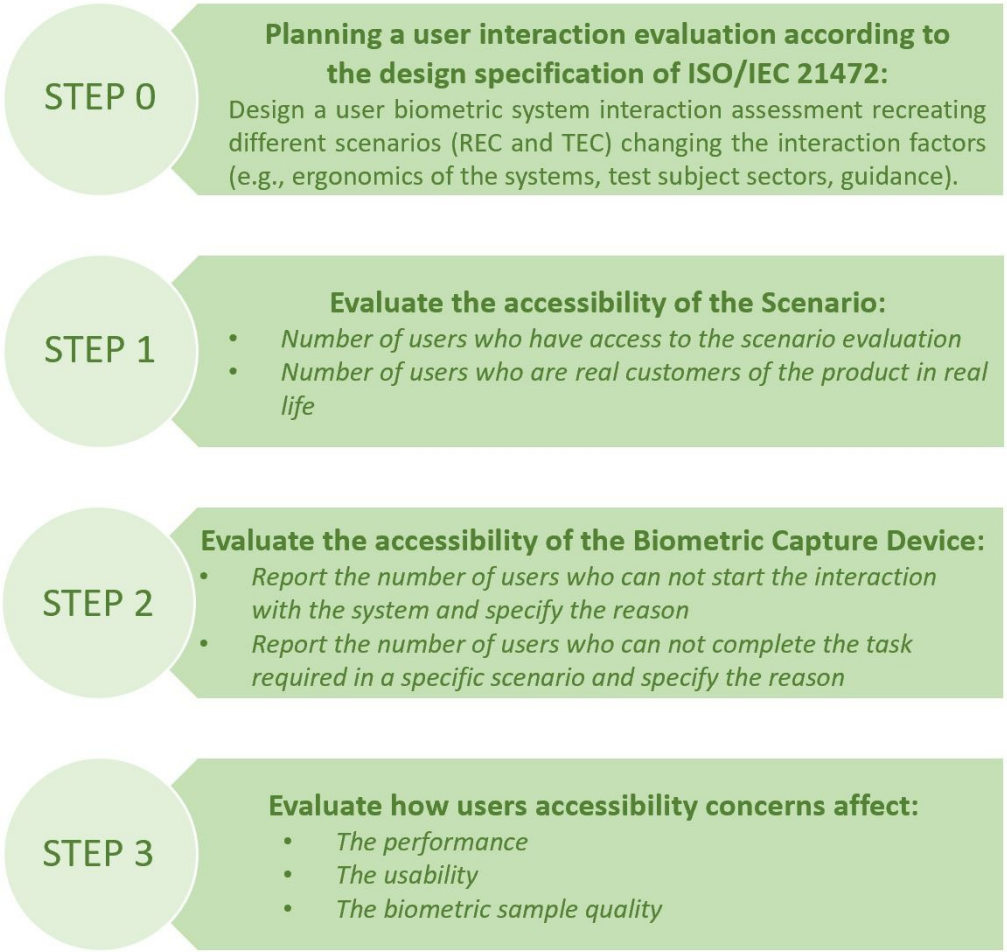


Figure 10: Accessibility Methodology.

Testing the accessibility of a biometric system means planning a user interaction evaluation and, then, evaluate whether the user had access to the recognition process. This means that the planning of the interaction evaluation is the starting point of the whole assessment process.

The methodology suggests designing a user evaluation according to the guidelines specified in the ISO/IEC 21472. This standard is taken as the basis (Step 0) for the application of the novel methodology because it specifies the nature of the factors that influence the user interaction (e.g. design of the system, test subjects, biometric modality, and training). Besides the influence on the system performance, these interaction factors are strictly related to the accessibility of the biometric recognition process.

Thus, as suggested in the ISO/IEC 21472, it is advisable to design more scenario evaluations to report the interaction factor influence on the accessibility of the system.

Once established the configuration of the user interaction evaluation, the methodology suggests achieving three steps to fully evaluate the accessibility of the biometric system. The specifications of these steps (and the relative accessibility metrics) are going to be detailed in the next subsections separately.

4.2.1 Step 1: Accessibility of the Scenario

Evaluating the accessibility of a scenario means to establish whether, or not, a given context is logistically accessible for all user categories. Through the accessibility of the scenario, we can establish whether the users can have access to a place in which it is required to interact with a biometric system. Thus, this aspect is related to the physical barriers that people may face in certain contexts, and that could make a biometric system unapproachable.

Another important information regarding the accessibility of the scenario is to observe how many data subjects were familiar with such a context of use. The aim is to observe if the participants of the evaluation already have access to a similar scenario in real life. An example of this is when assessing a mobile biometric system; in this case, is advisable

to report how many test volunteers are using this technology in their daily life. This is an aspect connected to the experience of the user which strongly influences the interaction with the system. Often, many people who are not used to interact with the new technologies (especially mobile devices) refuse to use them.

Considering these two aspects, the following metrics are proposed to evaluate the accessibility of the scenario:

• Number of users who have accessed the specific scenario.

• Number of users who are consumers of the product in real life.

The accessibility of the scenario is not strictly related to biometrics, although it is important to understand which are the causes that prevent people from using a biometric solution.

4.2.2 Step 2: Accessibility of the Biometric System

This part is closely linked to the interaction between the user and the biometric recognition system. Several factors can prevent people from accessing the use of a system by interacting with the related biometric sensor. As explicitly specified in ISO / IEC 21472, the factors that influence the interaction between the user and the biometric system are various and each of them must be taken into consideration even when evaluating the accessibility of a biometric system. A certain configuration, design, or typology of the biometric sensor can make the interaction with a recognition system impossible. A user with low hand dexterity may not interact with fingerprint sensors. Users with mobility concerns affecting their body posture could not correctly look at the camera for face recognition purposes. At the same time, people with cognitive or memory issues may find many difficulties in repeating several times a biometric recognition procedure.

Thus, to establish the accessibility of a biometric system is recommendable to report:

- the number of users who cannot start the interaction with the system and the reason.
- the number of users who cannot complete the task required in a specific scenario and the reason.

The numbers of users that can start or complete a recognition process were already used as accessibility metrics in [62]. The formal methodology advises reporting, besides these numbers, the causes that prevent users from starting the interaction with a biometric system or from finishing a task required in a recognition scenario.

4.2.3 Step 3: How the accessibility concerns impact on the outcome of the biometric process

In this part, the accessibility is referred to the health status of the user. Motor and cognitive concerns influence how a user interacts with a specific system. Even if a user attains to interact with a system to complete a biometric recognition process, his accessibility concerns may have an important effect on the outcome of the system. For this reason, it is equally important to report how the accessibility concerns of the users affect the outcome of the system in terms of:

- The usability
- The sample quality
- The performance

To understand how the accessibility issues can affect these aspects, it is necessary to enrol in the user interaction evaluation of different groups of data subjects. Thus, the data crew must gather a control group (people with no motor or cognitive issues) and, besides, an accessibility group (users affect by motor or cognitive pathologies).

By comparing the results obtained from the different groups, it will be possible to report the extent to which each accessibility characteristic of the user affects the outcome of the biometric system.

Chapter 5 Set-up of the Evaluations

This Chapter contains the details regarding the set-up of the two evaluations carried out to validate the Methodology proposed in Chapter 4.

Considering that the users provided their sensitive information, it is worthwhile to start the chapter by discussing the ethical implications behind the data collections.

Furthermore, each evaluation set-up is described separately presenting the biometric systems, the devices, the data crews, and the workflow of the experiment.

5.1 Ethical Implications

The European General Data Protection Regulation (GDPR) [63] specifies how to collect sensitive data from the citizens guaranteeing their security and privacy. The regulation categorizes personal, demographic, health, and biometric data as sensitive information.

Thus, before starting each experiment every user was required to sign a Consent (Annex 1) and Information Document (Annex 2). Through this document, participants

were informed regarding the goals of the evaluation and how their sensitive data would be stored and processed. The demographic, biometric, and health (e.g., accessibility concerns) information of the users was collected and stored in an encrypted database to which just the evaluation operators had access. Sensitive data were stored in the database for a period of one year after the beginning of the evaluation. During this period time, users can exercise their right to request the data holders to delete their saved data from the database.

5.2 Access Control System through Biometric Recognition

For the first experiment, a multi-modal biometric system was developed to be used in an access control context. The system allowed the data subjects to open a door after performing a fingerprint or a face recognition process employing different biometric sensors. According to the ISO/IEC 21472 [5], it is a must to change the evaluation conditions during the user interaction assessment. Hence, the system was tested by the users in 4 different scenarios (Figure 11).

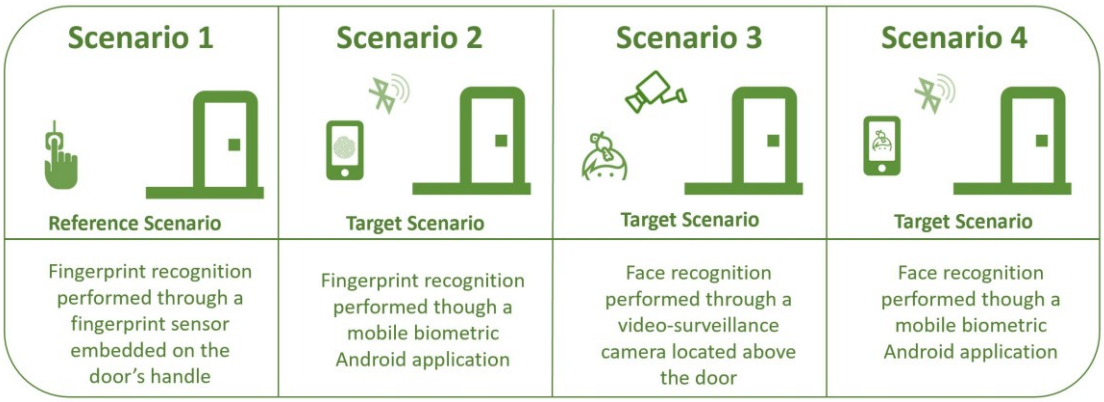


Figure 11: Diagram of the different scenarios of the first experiment.

The differences between the 4 scenarios, precisely, reside in "*the design, the position, and the condition of the biometric system and/or its biometric capture device*" [5]. In fact,

in the first scenario (that can be considered as the baseline scenario under reference evaluation conditions), users presented their fingerprint traits to a biometric scanner embedded on the door's handle.

In the third scenario, participants had to look to a camera above the door for providing their face parameters for the recognition process.

During the second and fourth scenarios, users were provided with a smartphone in which there was pre-installed an Android Application specifically developed for the experiment. Thanks to the app, participants could perform the fingerprint authentication (Scenario 2) and face recognition (Scenario 4) processes required to access the door in those parts of the evaluation.

Besides the validation of the formal methodology, the system was designed with this configuration for several reasons:

- To eradicate the accessibility barriers that people with mobility and cognitive issues may face in daily access control scenarios. This system was specially developed for all those users that have a problem when approaching the door or in turning the key in the lock (e.g., elderly customers, people in wheelchairs, and people with cognitive issues).
- To analyse and compare the user interaction when participants perform a recognition using a traditional biometric sensor (scenarios 1 and 3) and when they are using mobile biometric applications (scenarios 2 and 4).
- To establish which one is the most suitable biometric recognition solution for each category of user (e.g., young people, elderly user, and accessibility groups) in access control scenarios.

5.2.1 Devices and Applications Used During the Experiment

The door place of the access control system was recreated in a laboratory environment. Additionally, different smart devices were used and three applications (1 C# and 2

Android) were developed to enable the users to complete the tasks required in each scenario.

Firstly, the door was simulated using a Sony Xperia Tablet Z through a specific Android app. Every time that users completed the recognition tasks in each scenario, the application showed a door opening in the output screen.

For the first scenario, users presented their fingerprint traits to an EikonTouch710 [64] fingerprint sensor. This fingerprint scanner is a capacitive sensor (size: 12.8 x 18.00 mm) able to capture fingerprint samples as 256 greyscale images with a resolution of 508 ppi (pixels per inch). While the face images in the third scenario were captured through an Internet Protocol (IP) camera: the AXIS M1011 Network Camera [65]. This device is widely applied in real contexts as a video-surveillance device. The EikonTouch710 and the IP camera was connected to a computer to store the image of fingerprint and face of each one of the participants. The storage of the biometric data was carried out through a C# application. During the data acquisition, this application was managed by an operator (Figure 12).

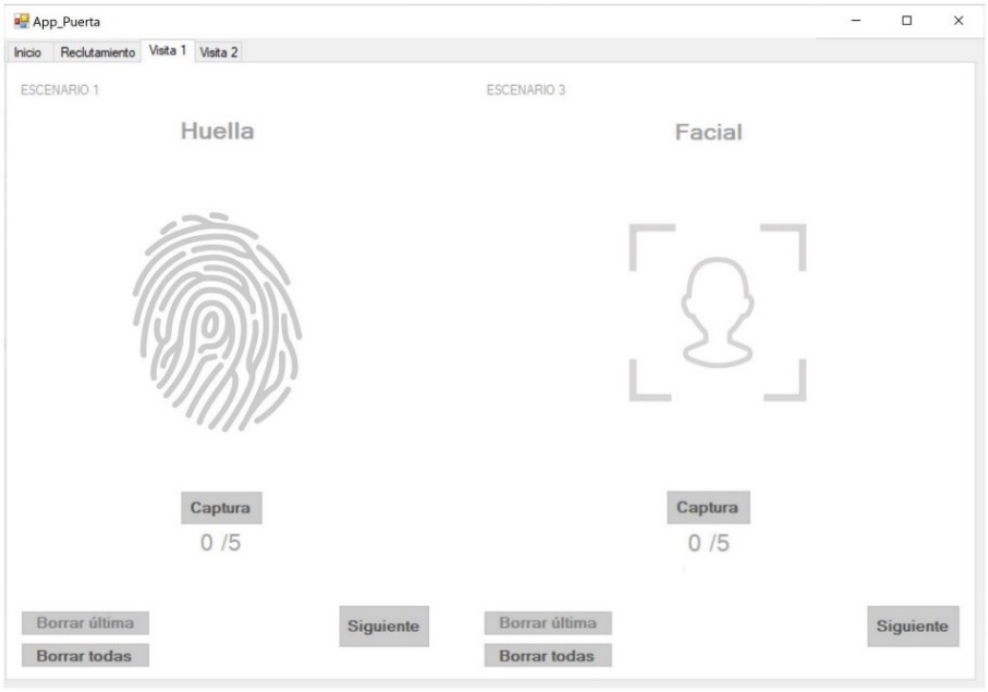


Figure 12: Screenshot of the C# application implemented for collecting the biometric samples throughout the scenario 1 and 3.

During the second and the fourth scenario, the data subjects were provided with a smartphone OnePlus 3T [66] (5,5" screen and 152,7 x 74,7 x 7,35 mm of size). In this device, a specific application (Figure 13) was installed to support the user during the smartphone scenarios.

The fingerprint recognition was carried out through the Android fingerprint security tool which does not allow the extraction of the biometric samples. Thus, no fingerprint sample could be collected in the second scenario.

The face recognition activities of the application were developed by implementing the OpenCV [67] library for face detection, storing in the smartphone's background the selfie photos taken by the users in the fourth scenario.

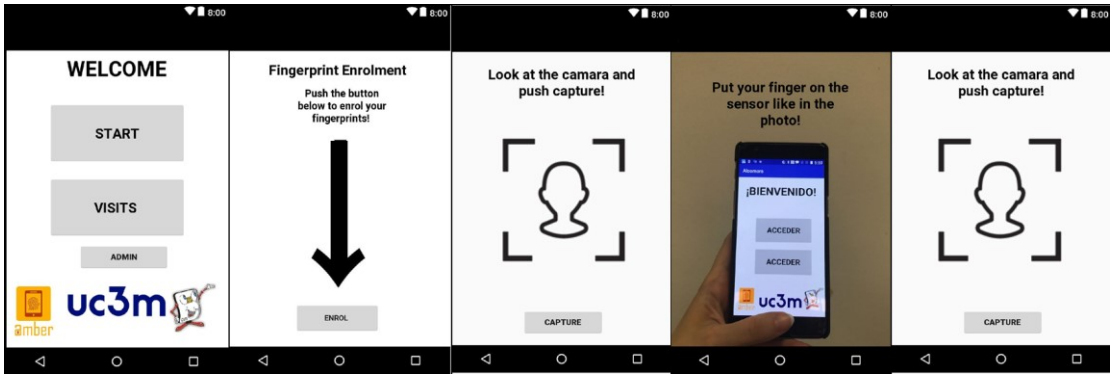


Figure 13: Android App's screenshots; a) starting interface of the application, b) fingerprint enrolment interface, c) face enrolment interface, d) fingerprint verification interface, and e) face recognition interface.

In all scenarios, the recognition process ended successfully every time the biometric sensors or the smartphone application recognized the users' faces or fingerprints. When this happened, a Bluetooth signal was sent from the C# (scenario 1 and 2) or the Android app (Scenario 3 and 4) to the Tablet showing to the users a door's opening.

5.2.2 Test Subjects

A group of 48 volunteers was recruited to interact with the proposed control system. The data crew was divided into two subgroups: control and accessibility.

The first group, the control group, gathered 31 users without any cognitive or mobility concerns (Table 5).

Table 5: Demographical information of the Control group.

Gender	Female	20
	Male	11
Age	Later adolescence (18 – 25 y/o)	16
	Early adulthood (26 – 30 y/o)	4
	Middle adulthood (31-50 y/o)	5
	Later adulthood (50 y/o and up)	6
Instruction level	No Instruction	1
	High School	14
	Bachelor's Degree	14
	Master's Degree	2
Experience with Technology	Computers	27
	Smartphones	30
	Tablets	23
	Biometrics Sensors	31
	Mobile Biometrics	21

The other 17 users composed the accessibility subgroup and presented different mobility and cognitive concerns. The accessibility group was recruited thanks to the

collaboration with local centres for rehabilitation and support for people with mobility and cognitive problems.

Table 6 shows the information regarding demographics and accessibility of these data subjects.

Table 6: Demographic and health data of the Accessibility group.

Gender	Female	9
	Male	8
Age	Later adolescence (18 – 25 y/o)	1
	Early adulthood (26 – 30 y/o)	5
	Middle adulthood (31-50 y/o)	6
	Later adulthood (50 y/o and up)	5
Instruction level	No Instruction	15
	High School	2
	Bachelor's Degree	1
Experience with Technology	Computers	3
	Smartphones	5
	Tablets	3
	Biometrics Sensors	17
	Mobile Biometrics	2
Accessibility problems	Developmental Issues	6
	Learning Issues	9
	Motor Issues	7

5.2.3 Evaluation Workflow

The entire evaluation workflow is shown in the image below (Figure 14).

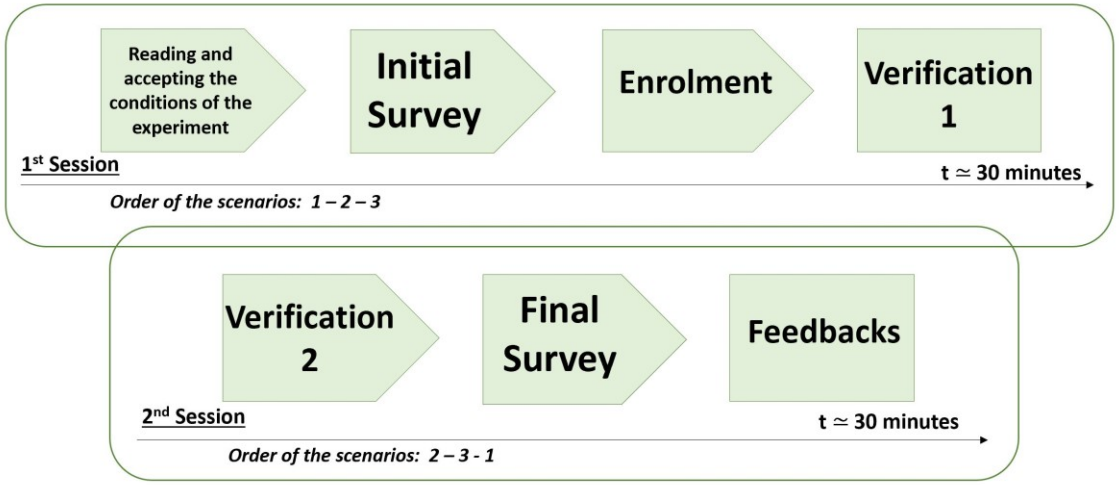


Figure 14: Workflow of the first experiment.

The experiment was divided into two different sessions: the first and second visits. These two parts were split into two different weeks. The 7-days delay between the first part and the second part of the experiment is imposed to guarantee that the users do not get accustomed to the system.

At the begging of the first session, the participants signed the Consent and Information Documents forms that contained all the information regarding the collection and the storage of their sensitive data (personal, health, and biometric information). After that, they filled an initial questionnaire regarding their demographical data and their previous experience with technology and biometric recognition devices.

Volunteers started to interact with the biometric system during the enrolment. In this phase, the data crew was asked to provide their biometric traits to the biometric sensors to store the template of their biometric traits (Figure 15).

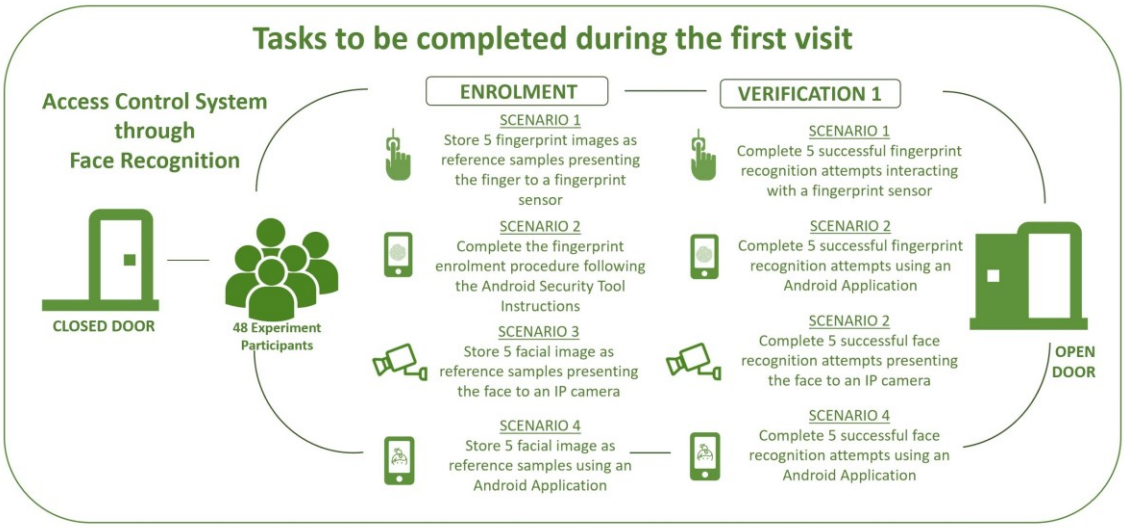


Figure 15: Tasks completed during the first visit.

Scenario 1, 3, and 4 required the Volunteers to interact with the biometric system during the enrolment. In this phase, the data crew was asked to provide their biometric traits to the biometric sensors to store the template of their biometric traits.

Scenario 1, 3, and 4 required the user to store five samples of their fingerprint (scenario 1) and face (scenarios 3 and 4) interacting with the biometric sensor. The enrolment of scenario 2 was carried out through the Android Security Setting tool. Regarding the fingerprint samples, participants stored their index finger.

Later, during the first verification, the participants were asked to complete 5 verification attempts for each scenario.

The second verification consisted of the same tasks as the ones on the first verification however, the order of the evaluating scenarios was the inverse of the first one (this time the following order was 4-3-2-1 instead of 1-2-3-4). This change was established to make the interaction more dynamic and to not let the participant get used to the system (Figure 16).

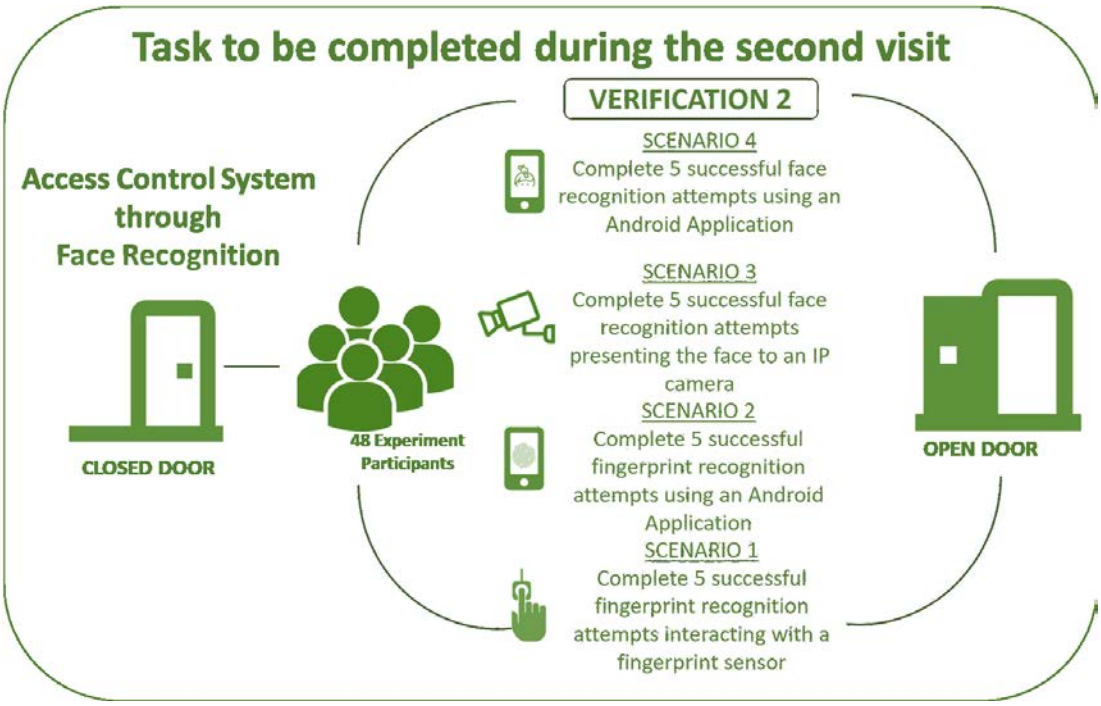


Figure 16: Tasks that were completed during the second visit.

Once the users ended with this part, they were asked to fill the final questionnaire to provide their opinion regarding the comfort, security, and easiness of the system and their feedbacks about the interaction with the biometric sensors.

5.3 Mobile Fingerprint Authentication System for Retail Payments

Nowadays, there are hundreds and hundreds of banking apps that support customers during retails payments. To complete money transactions, these applications establish connections between the smartphone and the sales point once authenticated the biometric traits of the users. The connection is generally based on NFC (Near Field Communication) technology that can reach just a few centimetres in communications. Thus, users are forced to approach their smartphones to the checkout to finalize the money transaction.

For the second experiment, a mobile biometric system for retail payment was developed in order to establish Bluetooth connections with the checkouts.

Bluetooth signals can reach up to 10 metres of distance. This means that through this system the participants could complete the retail payments without having to approach the cashier. Thus, this mobile biometric system eradicated all the accessibility inconveniences that arise when paying at supermarket stores especially in the cases of customers affected by mobility concerns.

The mobile biometric system consisted of an Android payment application based on fingerprint authentication. When people arrived at the supermarket's checkout and the cashier completed the bill receipt, the PoS (Point of Sales) terminal sent a payment request to the customer's smartphone. By accepting the payment notification, users were asked to provide their fingerprint to the smartphone's sensors. The payment transaction ended when the biometric authentication of the customer was completed.

Through this evaluation the main purpose was to establish whether the application of Biometrics could improve the usability and accessibility in payment scenarios. According to the ISO/IEC 21472, the application was tested in three scenarios using smartphones (Figure 17) with the fingerprint sensors located on a specific side of the device: front, lateral, backside.

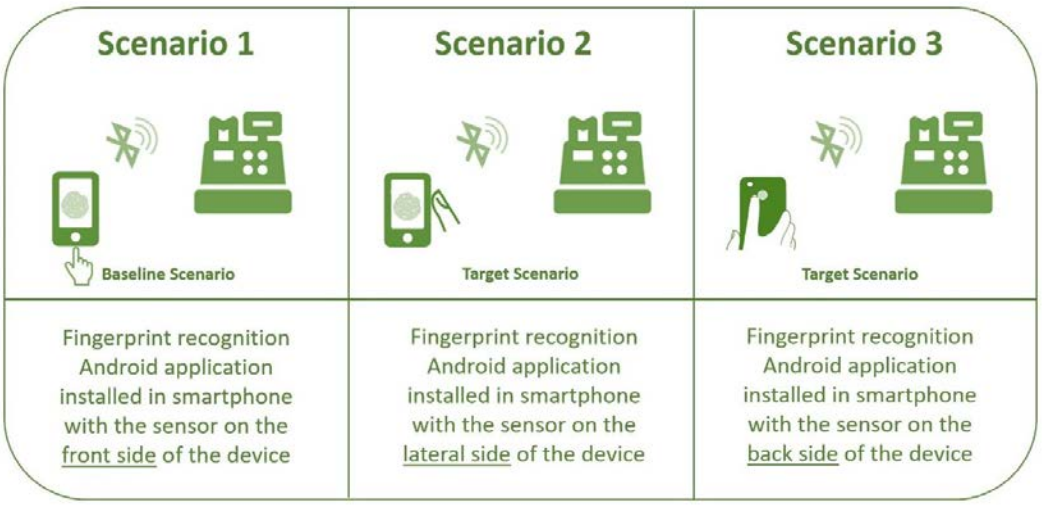
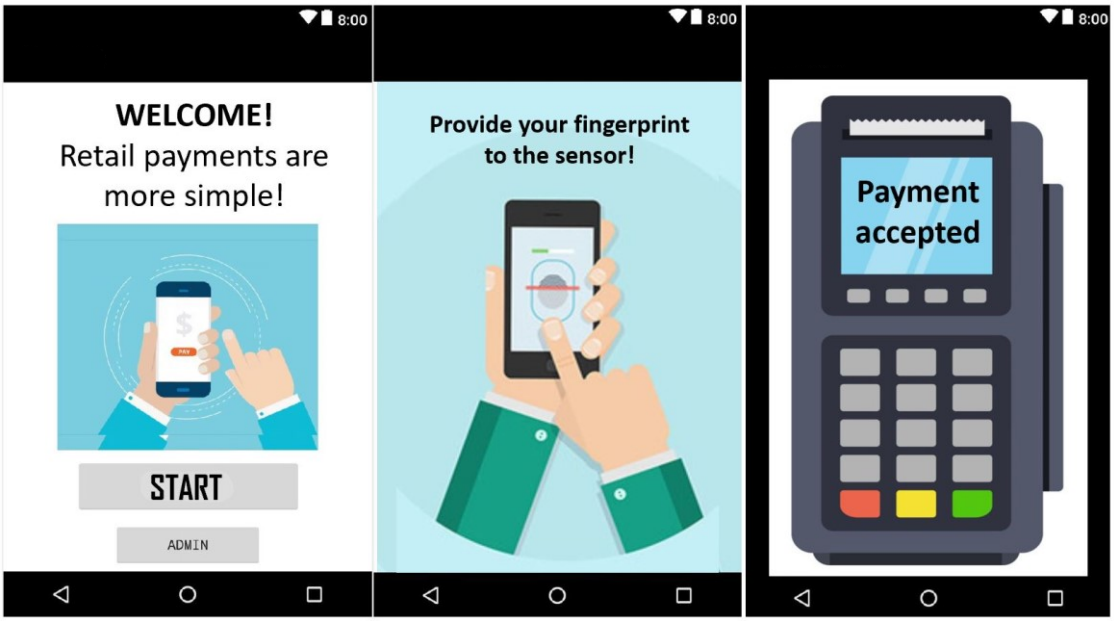


Figure 17: Different scenarios of the second user interaction evaluation.

5.3.1 Devices and Application Used During the Experiment

For this experiment, it was developed a fictitious Android banking app that allowed users to complete retail payment transactions. The payment process carried out through the mobile application can be seen in Figure 18.



a) b) c)

Figure 18: Interface of the application screen at different payment stages; a) welcome interface, b) fingerprint authentication interface, and c) notification of payment's acceptance.

Once the data subjects completed the fingerprint enrolment through the Android security settings, they could access the application to conclude the money transaction (Figure 18.a). Hence, the application asked them to present their fingerprint traits to the biometric sensor (Figure 18.b). When the recognition process ended successfully the application notified that the payment was accepted (Figure 18.c).

During the evaluation, participants were asked to interact with three different mobile devices. As already mentioned, in each smartphone, the fingerprint sensor was located on a different side of the device.

The smartphone characteristics are illustrated in the following table (Table 7)

Table 7: Morphology and characteristics of the employed smartphones.

Device	Acronym	Fingerprint Sensor Position	Sensor Shape	Device Size	Screen Size
OnePlus 3T [66]	D1	Frontal side	Rectangular	152,7x74,7x7,4 mm	5,5”
Sony Xperia XZ [68]	D2	Lateral Side	Rectangular	146x72x8,1 mm	5,2”
Neffos C9 [69]	D3	Back Side	Circular	158,7x76,6x8,5 mm	5,9”

The supermarket check-out was simulated with another Android application running on a Sony Xperia Tablet Z.

Each time a participant performed a successful recognition attempt, the smartphone sent a Bluetooth signal to the tablet device to notify the result of the transaction.

5.3.2 Test Subjects

A group of 21 volunteers was recruited to evaluate the mobile payment system. As the first experiment, even this data crew was divided into two subgroups: control and accessibility.

The control group gathered 9 volunteers without any cognitive or mobility concerns. Table 8 encloses their demographic and experience information.

Table 8: Demographic experience information of the control group.

Gender	Female	4
	Male	5
Age	Later adolescence (18 – 25 y/o)	1
	Early adulthood (26 – 30 y/o)	2
	Middle adulthood (31-50 y/o)	3
	Later adulthood (50 y/o and up)	3
Instruction level	High School	5
	Bachelor's Degree	1
	Master's Degree	3
Experience with Technology	Computers	9
	Smartphones	9
	Tablets	8
	Biometrics Sensors	9
	Mobile Biometrics	5

The other 15 users belonging to the accessibility group were recruited thanks to the collaboration of local centres for rehabilitation and support for people with mobility and cognitive problems. Table 9 provides information regarding their demographic data, their experience, and accessibility concerns.

Table 9: Demographic and health data of the accessibility group.

Gender	Female	6
	Male	9
Age	Middle adulthood (31-50 y/o)	10
	Later adulthood (50 y/o and up)	5
Instruction level	No Instruction	15

Experience with Technology	Computers	7
	Smartphones	10
	Biometrics Sensors	16
	Mobile Biometrics	0
Accessibility problems	Leg Concerns	2
	Hand Concerns	3
	Cognitive Concerns	15

5.3.3 Evaluation Workflow

The workflow of this evaluation was similar to the first one. Even in this case, the experiment was split into two different sessions: first and second visit, 7 days apart.

The whole process workflow can be seen in Figure 19:

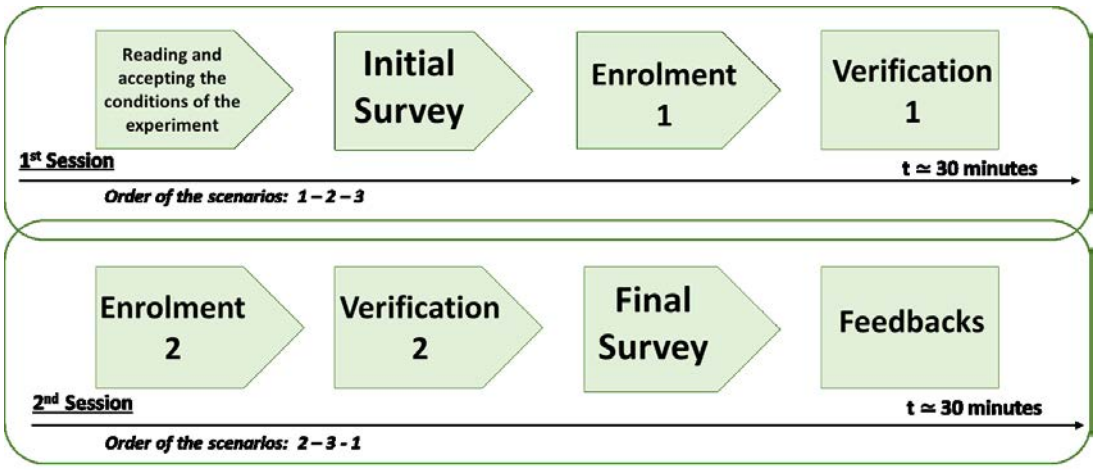


Figure 19: Workflow of the process.

Each participant read and signed the Consent and Information form before taking part in the evaluation. Their demographical data and their previous experience with technologies and biometrics were collected through an initial survey that all users were asked to fill.

Later, the participants started the first visit (Figure 20) enrolling their fingerprint traits, belonging to their handedness index finger, in all through the Android security setting tool using each one of the three devices provided them. During the first verification, volunteers were asked to complete 5 verification attempts interacting with the D1, D2, and D3.

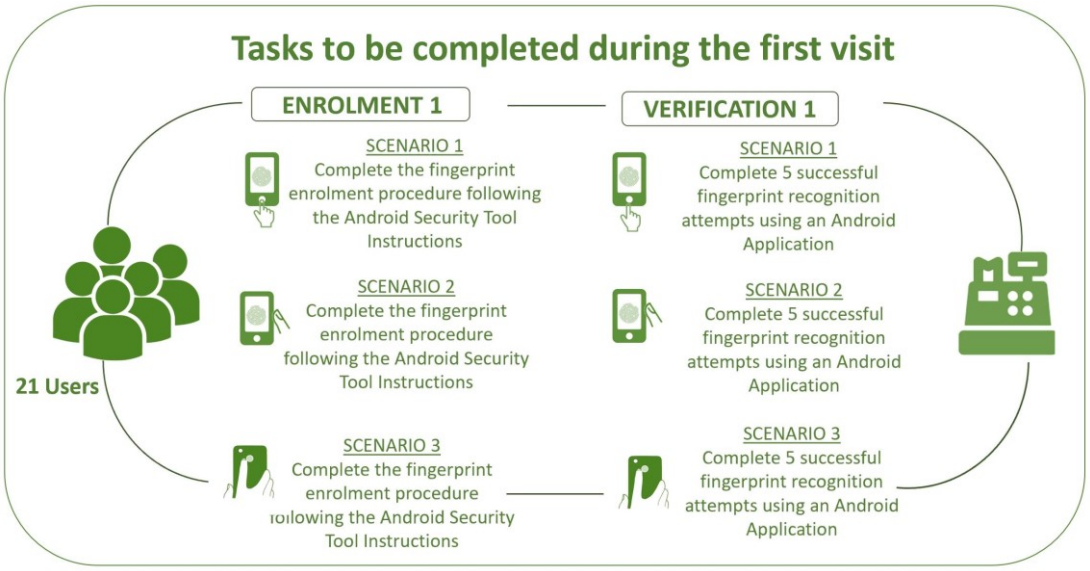


Figure 20: Diagram of the tasks that are completed during the first visit.

One-week later participants were asked to complete a second visit (Figure 21).

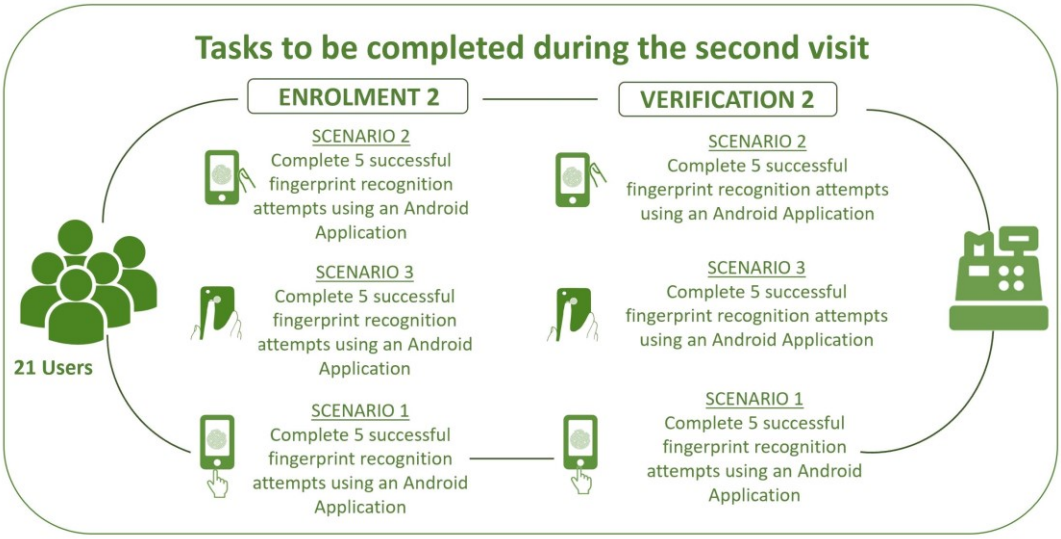


Figure 21: Diagram of the tasks that are completed during the second visit.

Since the Android Fingerprint tool allowed the storage of just 5 fingerprint templates, users were required to perform another enrolment through the 3 mobile devices. Later, participants completed the second verification (5 identification attempts) this time with another smartphone interaction order: D2, D3, and D1.

The experiment ended with the last task required to the users: fulfilling a final survey to collect their opinions regarding the interaction with the system and the mobile devices.

Chapter 6 Access Control System through Biometric Recognition

In this chapter, we are going to report the results obtained evaluating the first system we developed: an access control system based on biometric recognition.

As previously mentioned, the system was designed to test the application of fingerprint and face recognition in daily scenarios (such as opening a home's door). Besides, the experiment also aimed to compare the interaction between mobile and not-mobile biometric sensors.

Thus, participants who took place in the experiment completed four different scenarios interacting with different biometric sensors (mobile and not-mobile based) as shown in figure 22.

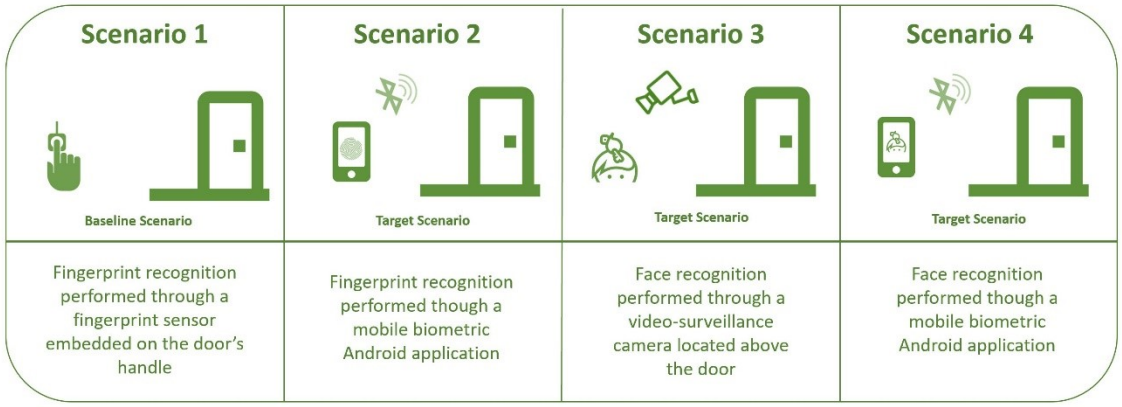


Figure 22: Scenarios considered for evaluation of the Biometric Access Control System.

The data collected during the experiment was processed and evaluated according to the novel methodology proposed in Chapter 4. Hence, the results will be discussed by analysing each step of the methodology separately.

The Chapter ends with a brief discussion about the main findings (related to the accessibility of the system) reached examining the results obtained.

6.1 Result Analysis according to the Novel Methodology

The novel methodology, that we proposed to report the accessibility of biometric systems, suggests conducting an evaluation articulated in three steps:

- to evaluate the extent to which users have access to the evaluation scenario.
- to evaluate the extent to which a biometric system is accessible for the users and how easily they complete a recognition process.
- to study how the accessibility concerns of the users affect the whole outcome of the recognition process (in terms of performance, usability, and sample quality).

Before analysing each one of these points alongside the next subsections, we are going to report useful information regarding the characteristics of volunteers who took part in the experiment as data subjects.

6.1.1 Data Crew

48 volunteers joined the experiments as data subjects, and they were split into two main groups: control and accessibility group.

The control group was composed of 31 users without any kind of accessibility issues. According to each participant's age, this group was split into 4 subgroups: Age 1, Age 2, Age 3, and Age 4.

While the accessibility group gathered 17 users affected by different types of accessibility issues. This group was divided into three smaller subgroups according to the participant accessibility concerns: developmental, learning, and motor issues.

The main characteristics of each user group are reported in the table below (Table 10):

Table 10: Main characteristics of the experiment's participants.

Group	Subgroup	Characteristics	Number of Users
Control	Age 1	Later adolescence (18 – 25 y/o)	16
	Age 2	Early adulthood (26 – 30 y/o)	4
	Age 3	Middle adulthood (31-50 y/o)	5
	Age 4	Later adulthood (50 y/o and up)	6
Accessibility	Developmental	6 users with congenital related disorders affecting their cognitive abilities	6
	Learning	9 users with intellectual disorders affecting their attention abilities	9
	Motor	Temporary Wrist Issues (1 user) Leg Issues on a wheelchair (2 users) Leg Issues using crutch (1 user) Arm Issues (2 users) Hand issues (1 user finger amputated)	7

6.1.2 Accessibility of the Scenario

This subsection deals with the accessibility of the scenario. This first step of the novel methodology aims to analyse if the proposed system is logistically accessible to the entire data crew.

All the users, belong to both control and accessibility groups, were able to access the scenario recreated to evaluate the access control system. Besides, table 11 reports the information regarding the previous experience of the user with smartphones and mobile biometrics.

Table 11: Users' experience with smartphones and mobile biometrics.

Subgroup	Smartphone Owners	Mobile biometric experience
Age 1	16/16	14/16
Age 2	4/4	3/4
Age 3	5/5	3/5
Age 4	5/6	0/6
Developmental	0/6	0/6
Learning	2/9	0/9
Motor	3/7	1/7

Regarding the control groups, almost all the younger users owned smartphones. Additionally, most of them (14 over 16) were using mobile biometric tools to access and unlock their smartphones or to access banking apps. While elderly users, gathered in the Age 4 group, declared no experience with mobile biometrics (one of them didn't even have a smartphone).

Different situations happened for the accessibility groups. Even if, volunteers belonging to the developmental, learning, and motor issues group could assess the scenario, just a few participants were also smartphone users. Additionally, only one

participant (the user affected by temporary wrist issues) already experienced the interaction with mobile biometric applications.

6.1.3 Accessibility of the system

This section deals with the accessibility of the system. Thus, the following table (table 12) reports the number of users that cannot start interacting with the system in each scenario.

Table 12: Number of users who cannot start the interaction with the system.

Subgroup	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Age 1	0	0	0	0
Age 2	0	0	0	0
Age 3	0	0	0	0
Age 4	0	0	0	0
Developmental	1	1	2	0
Learning	3	2	2	1
Motor	1	3	0	0

Users coming from the control subgroups were able to start the interaction with the system in each scenario.

Regarding the accessibility groups, one user with developmental issues was not able to interact with the fingerprint recognition in scenario 1 and 2, and 2 users can not start the face recognition process through the IP camera (it was very difficult for them to understand how to look at the camera).

Learning issues affected several users in understanding how to interact with the biometric sensors and with the mobile application. While motor concerns impeded 3 users to start the Android enrolment for the fingerprint recognition in scenario 2 (they had issues in holding the device and provide their finger to the button sensor).

In table 13, the numbers of users who cannot complete any part of the evaluation are reported.

Table 13: Number of users who could not complete the interaction with the system.

Subgroup	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Age 1	0	0	0	0
Age 2	0	0	0	0
Age 3	0	0	0	0
Age 4	1	0	1	0
Developmental	1	2	4	0
Learning	2	2	2	2
Motor	1	2	3	2

Age 1, Age 2, and Age 3 subgroups completed the tasks required alongside the 4 scenarios. While a user belonging to Age 4 did not end scenario 1 and scenario 3 due to the tiredness at the end of the second visit.

Several users with developmental and learning issues had many difficulties in remembering how to interact with the sensors for completing the task required them. Thus, some of them did not conclude the whole experiment.

Besides, some users affected by motor issues were not able to complete the tasks required in the proposed scenarios due to a lack of hand dexterity (in fingerprint

authentication scenarios) and due to motor issues that impeded them to have a right posture (during face recognition scenarios).

6.1.4 How the accessibility concerns impact on the outcome of the biometric process

Across the next subsections, we are going to analyse the impact of age and accessibility issues on the biometric recognition process. Thus, we are going to report the results obtained by evaluating the performance, usability, and biometric sample quality considering each data subject subgroups separately.

6.1.4.1 Performance

This part is focused on evaluating the performance of the recognition processes completed by the user in all the 4 scenarios of the experiment. The results are gathered according to the scenario.

6.1.4.1.1 Performance of Scenario 1

In the first scenario, the user interacted with an external fingerprint sensor. Thus, we were able to extract the biometric samples creating a specific fingerprint image database.

Once the experiment ended, we processed the database using VeriFinger SDK [70] a commercial software produced by Neurotechnology [71].

The software compared all images captured during the verifications with all the fingerprint templates stored during the initial phase of the experiment (enrolment). The comparison is based on the analysis of the minutiae presented in each fingerprint image sample. The result of this comparison is a matching score indicating the degree of similarity between a fingerprint sample and each image stored in the database as

templates. Through the matching scores, we could calculate the false negative and the false positive rate. Thus, we report in table 14 the EER of the fingerprint recognition relative to the first scenario.

Table 14: EER scores for scenario 1

Subgroup	Visit 1	Visit 2
Age 1	0,5739%	0,3592%
Age 2	0%	0%
Age 3	0.071%	0,502%
Age 4	0,5319%	6,42%
Developmental	0,2415%	0,7246%
Learning	2,1739%	1,2077%
Motor	3,15%	0,7%

Between the first and the second visit, the EER decreased (meaning an increment of the performance) for all the groups except the Age 4 and the Developmental Issues subgroups. This can be addressed to the tiredness of the elderly users and the participant with developmental concerns at the end of the second visit (when the scenario 1 was evaluated as the last one).

Besides, the highest percentages of EER were recorded for the Learning and Motor issues subgroups. This was due to their accessibility concerns: users affected by Learning issues had some interaction problem in collocating the fingerprint on the sensor; while Motor issues affected that mobility of the participants belonging to the last subgroup.

6.1.4.1.2 Performance of Scenario 2

During the second scenario, we asked participants to perform a fingerprint recognition process using, this time, a specific Android biometric application. Since the Android security tool does not allow the extraction of the fingerprint images, the performance of the second scenario was analysed reporting the percentage of successful recognition attempts. Although the sample extraction is blocked, the mobile application we developed could register the results of each verification attempts performed by the participants.

Thus, the performance is analysed through the percentage of successful recognition attempts during each verification phase.

Regarding the control population (Figure 23), the percentage of successful recognition attempts increased between the fist and the second visit.

Younger groups (Age 1 and Age 2 subgroups) obtained high percentage of successful attempts during both visits. During the first verification, users belonging to the Age 3 and Age 4 did less successful attempts compared with the younger participants. Besides, the performance of the second visit for the older group is the lower among the control subgroups.

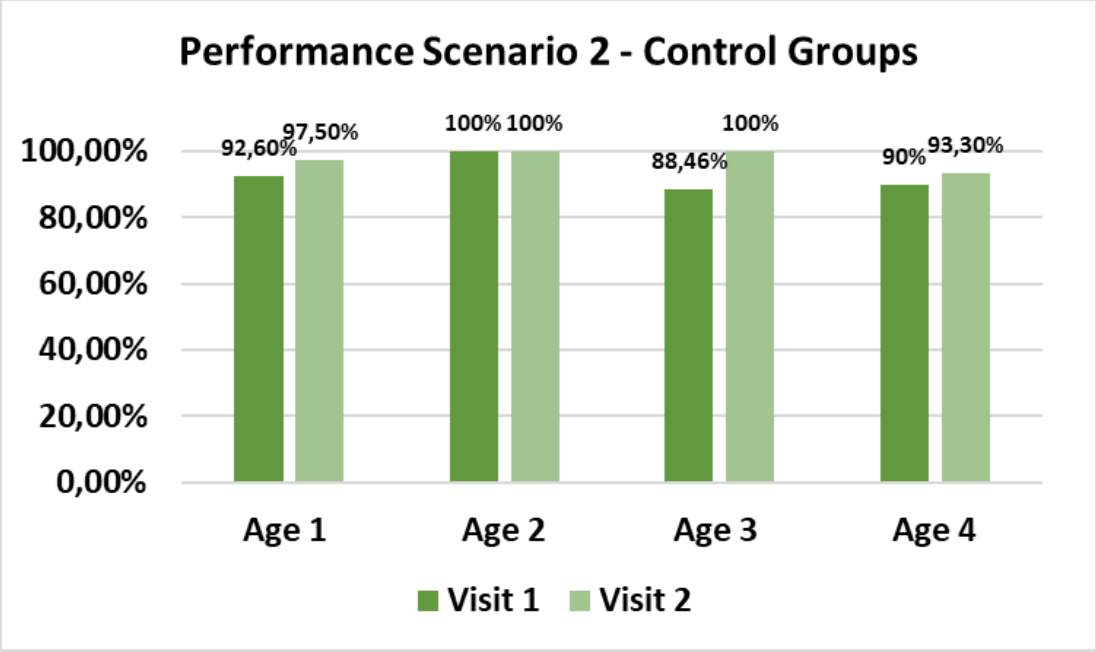


Figure 23: Percentage of successful recognition attempts performed by control populations.

The accessibility population also reported lower level of successful recognition attempts compared with the younger users. The lowest performance rates were obtained by the developmental issues subgroup followed by motor issues group. This because both groups found lots of interaction problems when presenting their fingerprint to the smartphone's sensor.

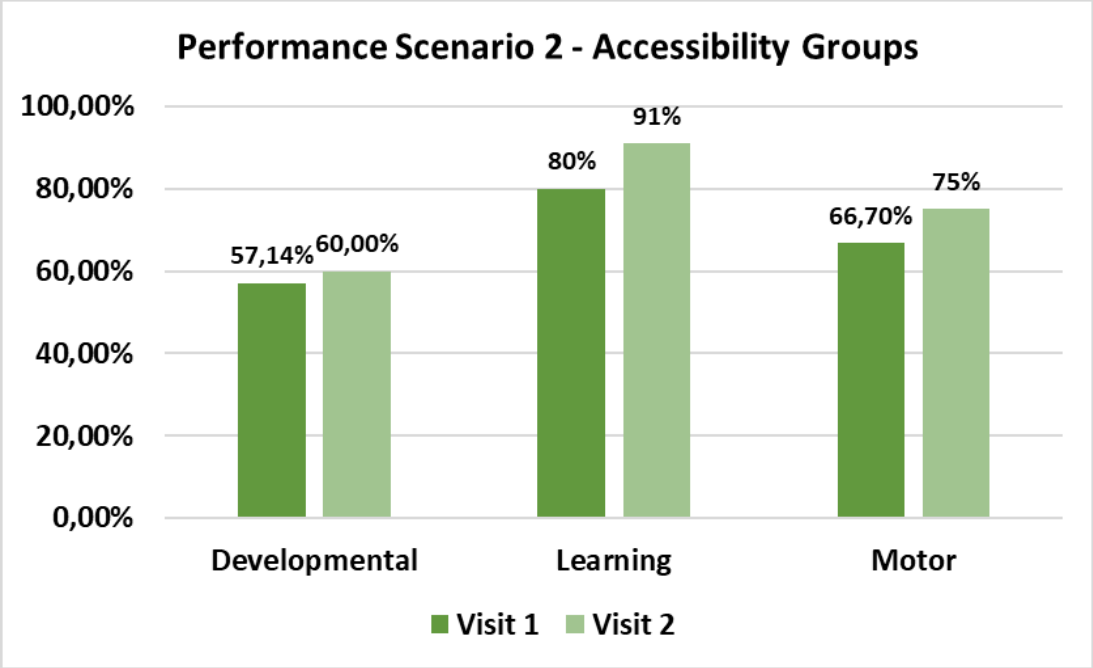


Figure 24: Percentage of successful recognition attempts performed by accessibility populations.

6.1.4.1.3 Performance of Scenario 3

In the third scenario, participants were asked to present their facial traits to an IP camera. The photo captured during this part of the experiment were stored in the face image database.

The performance of the face recognition processes carried out in Scenario 3 was calculated using another software released by Neurotechnology: VeriLook SDK [72]. This software works similarly to the VeriFinger SDK. Thus, each image captured during

the verifications were compared with the template database obtaining a matching score. Thus, even in this case, we extracted the EER scores (Table 15).

Table 15: EER scores for scenario 3.

Subgroup	Visit 1	Visit 2
Age 1	0,2588%	0,188%
Age 2	0,6056%	0,446%
Age 3	0,44%	0,478%
Age 4	0,8012%	0,96%
Developmental	2,313%	5,14%
Learning	1,2871%	2,34%
Motor	3,838%	4,1%

Younger users (Age 1 and Age 2 subgroups) gained experience in performing face recognition during scenario 3. This did not happen for the older users and the accessibility subgroups. As we can see the performance decreased between the first and the second visit.

6.1.4.1.4 Performance of Scenario 4

The images captured by the Android application were also stored in a specific facial image database for the scenario 4.

The facial photos were processed using VeriLook SDK. Thus, the performance analysis was been conducted with the same procedure applied in the third scenario (Figure 16).

Table 16: EER scores for scenario 4.

Subgroup	Visit 1	Visit 2
Age 1	0,2558%	0,6052%
Age 2	0,6056%	0,2438%
Age 3	0,4%	0,3581%
Age 4	0,103%	2,04%
Developmental	3,15797%	9,3%
Learning	3,33%	0,521%
Motor	4,7%	3,1%

Between the visits, the performance is almost constant for the control populations. While it decreases for the accessibility populations. The EER increased just for Age 4, probably due to the difficulty of the older users in remembering how to take selfie photos. Regarding the accessibility population, the EER scores are generally higher compared with the control populations demonstrating that Developmental, Learning, and Motor concerns affected the interaction between the participants and the mobile app.

6.1.4.2 Usability

The usability is going to be analysed according to the metrics specified in Chapter 4, this means reporting the efficiency, effectiveness, and satisfaction. The results are going to be accessed according to the experiment's phase and the group of users.

6.1.4.2.1 Efficiency

The efficiency was measured reporting the mean and the standard deviation of the seconds spent by different subgroups during the enrolment and the verifications. Due to the mobile app configuration and to the Android Security setting, it was not possible to store the enrolment time during scenario 2.

Regarding the enrolments (table 17), users belonging to Age 4 took more time to complete the enrolments of scenarios 1 and 3 compared with Age 1, Age 2, and Age 3. While during the enrolment of the fourth scenario, the younger users took more time. This was probably due to the younger users' experience practice in taking selfies for social networks. In fact, during this task, we observed that young users were more prone to find better poses and the right shot. Hence, this has lengthened the completion of the enrolment in scenario 4 for Age 1 and Age 2 subgroups.

On the other hand, the users affected by motor issues spent more time in completing the enrolments respect the developmental and learning issues. This was caused to the low level of hand dexterity of the volunteers belonging to the motor subgroup (three of them were affected by arm motor concerns).

Table 17: Mean (μ) and standard deviation (σ) of the time (in second) spent by users during the enrolments.

Subgroup	Scenario 1		Scenario 3		Scenario 4	
	μ	σ	μ	σ	μ	σ
Age 1	12,7	0,62	13,94	3,1	62, 13	19,7
Age 2	12	0,5	10	2,5	57,393	10.6
Age 3	12,6	0,7	13	3,77	83,7	32,4
Age 4	15,41	9,7	20,19	3,01	67,2	30,9
Developmental	19,5	12,23	19,13	7,9	31,7	2,2
Learning	17,78	14,91	18,52	5,94	30,6	17,88
Motor	21,27	10,26	20,67	3,03	40,7	17,6

When observing the efficiency of the verification visits, younger users completed faster the enrolment of scenarios 1 and 2 compared with the elderly participants (Table). While accessibility users spent more time completing the tasks in scenario 2 than the control subgroups (probably due to the lack of experience of the accessibility groups in using smartphones).

Completing the second visit of scenario 1 took more time compared with the first visits (table 18), this could be addressed to the user tiredness at the end of the second visit (when the scenario 1 was performed at the end of the session).

Table 18: Mean (μ) and standard deviation (σ) of the time (in second) spent by users during the visit 1 and visit 2 of scenario 1 and 2.

Subgroup	Scenario 1				Scenario 2			
	Visit 1		Visit 2		Visit 1		Visit 2	
	μ	σ	μ	σ	μ	σ	μ	σ
Age 1	2,9	1,4	3,58	1,44	3,6	2	3,1	1,1
Age 2	2,36	0,82	2,4	1,22	1,4	0,8	2,6	0,3
Age 3	3,3	2,03	8	1,71	3,2	0,8	3,7	1,4
Age 4	3,6	1,5	5,4	2,61	4,1	1,5	4,5	3,3
Developmental	2,9	0,61	3,18	1,07	5,18	1,91	3,3	2,9
Learning	2,3	1,01	4,2	3,22	5,47	1,9	5,46	2,87
Motor	2,9	0,64	3,4	2,4	4,48	2,1	3,14	2,5

Besides, during the first and second visits to scenario 3, younger users generally took less time than older users (Table 19).

Table 19: Mean (μ) and standard deviation (σ) of the time (in second) spent by users during the visit 1 and visit 2 of scenario 3 and 4.

Subgroup	Scenario 3				Scenario 4			
	Visit 1		Visit 2		Visit 1		Visit 2	
	μ	σ	μ	σ	μ	σ	μ	σ
Age 1	3,7	0,7	4,23	0,88	6,8	1,3	6,9	1,0
Age 2	5,18	2,02	4,31	0,82	5,7	1,2	6,9	1,3
Age 3	4,81	1,9	4,2	0,81	7,3	0,9	8	1,3
Age 4	4,5	0,9	3,8	0,54	7,5	1,2	8,6	2,3
Developmental	3,04	1,34	5,14	0,65	4,84	1,84	5,3	2,64
Learning	3,8	1,1	4,78	1,12	5,3	1,47	5,72	2,3
Motor	3,2	1,5	4,63	0,07	5,49	1,9	5,47	2,2

The accessibility groups took more time in completing the second visit compared with the efficiency of the first visit. This because several users belonging to these groups face lots of difficulties in remembering how to present the face to the IP camera. Besides, motor issues have prevented some users from having the correct pose for completing the face recognition process.

The times spent in completing the enrolments and verifications were evaluated also to demonstrate if the subgroups were statistically independent by conducting the ANOVA[73] test. From this analysis we obtained the following p-value scores: 0,0001 (Scenario 1), 0,0025 (Scenario 3), 0,00183 (Scenario 4) for the enrolments; 0,013 (Scenario 1), 0,02 (Scenario 2), 0,0122 (Scenario 3), 0,02113 (Scenario 4) for the visit 1; and, 0,05 (Scenario 1), 0,0203 (Scenario 2), 0,039 (Scenario 3),0,0247 (Scenario 4) for the visit 2. All the p-values are lower or equal to the null hypothesis value (0,005). This means that the data analysed are significant and the seven subgroups statistically independent.

6.1.4.2.1 Effectiveness

The effectiveness was evaluated through the percentage of incorrect interactions made by the user in the first and second verification. The incorrect interactions happened when users presented incorrectly interact with the biometric sensor. For instance, when users wrongly touched the fingerprint scanner (not pressing the biometric sensor area); or when participants mistook in using the mobile applications.

The percentage of incorrect interaction of scenario 1 and 2 are shown in the table 20. As the efficiency evaluation, even in this case was not possible to report the effectiveness data related to scenario 2.

The learning issues subgroup had several interaction problems in completing the task required alongside scenario 1. Among almost all the recruited subgroups, there is a quite general increment of incorrect interaction in the second visit confirming the users' tiredness at the end of the second visit.

Table 20: Percentage of incorrect interactions made by each population during the fingerprint recognition scenario (scenario 1 and 2)

Subgroup	Scenario 1			Scenario 2	
	Enrolment	Visit 1	Visit 2	Visit 1	Visit 2
Age 1	10%	16,25%	22,5%	11,25%	10%
Age 2	0%	5%	35%	0%	0%
Age 3	12%	8%	0%	4%	0%
Age 4	3,3%	13,3%	40%	6,67%	36,67%
Developmental	15%	10%	10%	5%	5%
Learning	14,3%	28,6%	42,85%	11,42%	17,14%
Motor	20%	10%	0%	0%	20%

Regarding the other two scenarios (table 21), elderly users (Age 4) made more mistakes when completing the tasks required in Scenario 3 and 4 (table) compared with

younger. We argue that this may be addressed to the lack of experience in using mobile biometric applications characterizing the older participants.

Besides, users affected by accessibility concerns (especially by developmental and learning issues) registered a high percentage of incorrect interactions during the second visit of scenario 4. Thus, cognitive concerns prevented the user from remembering how to successfully interact with mobile biometric applications.

Table 21: Percentage of incorrect interactions made by each population during the face recognition scenario (scenario 3 and 4).

Subgroup	Scenario 3			Scenario 4		
	Enrolment	Visit 1	Visit 2	Enrolment	Visit 1	Visit 2
Age 1	7,5%	16,25%	12,5%	3,75%	8,75%	20%
Age 2	5%	20%	0%	0%	5%	30%
Age 3	12%	8%	4%	0%	28%	48%
Age 4	6,67%	16,7%	23,3%	23,3%	20%	43,33%
Developmental	10%	10%	10%	25%	20%	50%
Learning	20%	31,42%	17,14%	0%	2,85%	31,42%
Motor	5%	10%	0%	0%	10%	25%

6.1.4.2.1 Satisfaction

Through the satisfaction we intended to analyse the users' opinions regarding the evaluation and the use of biometric in real access control scenarios. For this reason, we asked volunteers to fill a satisfaction survey at the end of the whole experiment.

Users were required to rate the comfort and the time needed to perform each scenario. The rates were expressed by scores between 1 and 5 (1 indicated the slowest and less

comfortable scenario, while 5 the faster and the most comfortable). Besides, they also expressed their preference for the scenario and the biometric trait.

In figure 23, there are shown the results of the satisfaction questionnaire fulfilled by the younger users (Age 1 subgroup). According to their opinions, scenario 2 (fingerprint recognition on a mobile device) was the more comfortable and the faster. While the less comfort and most slow was scenario 3 (face recognition with an IP camera).

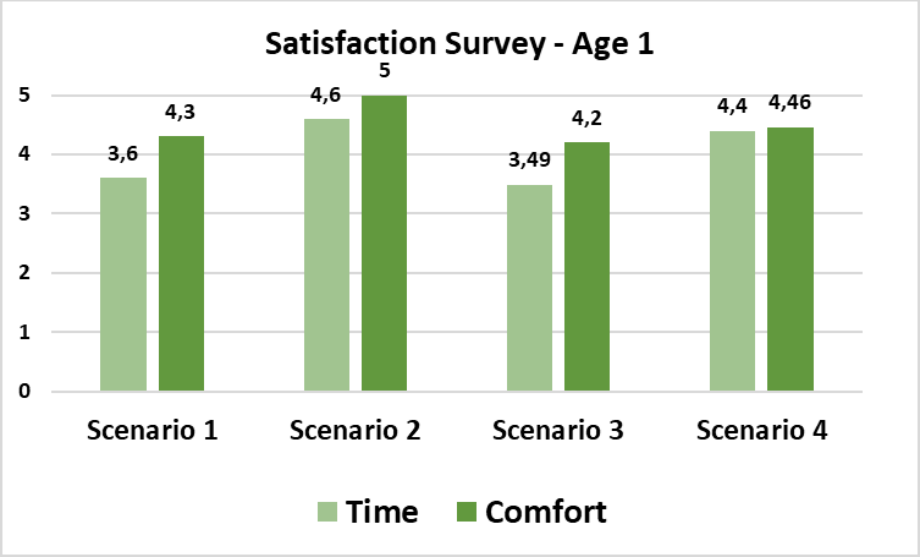


Figure 25: Satisfaction Survey results Age 1.

Besides, younger users also declared fingerprint recognition as the favourite authentication method in the access control scenarios. 89% of Age 1 group stated that they were willing to use mobile fingerprint applications to access their own home and even in other scenarios (e.g., banking).

Scenario 2 was rated as the best scenario in terms of time and comfort even by the Age 2 subgroup (Figure 24). While they rated scenario 4 as the slowest and less comfortable. Thus, also the majority of the Age 2 subgroup stated that they will prefer to use a mobile app based on fingerprint recognition for the access control scenario.

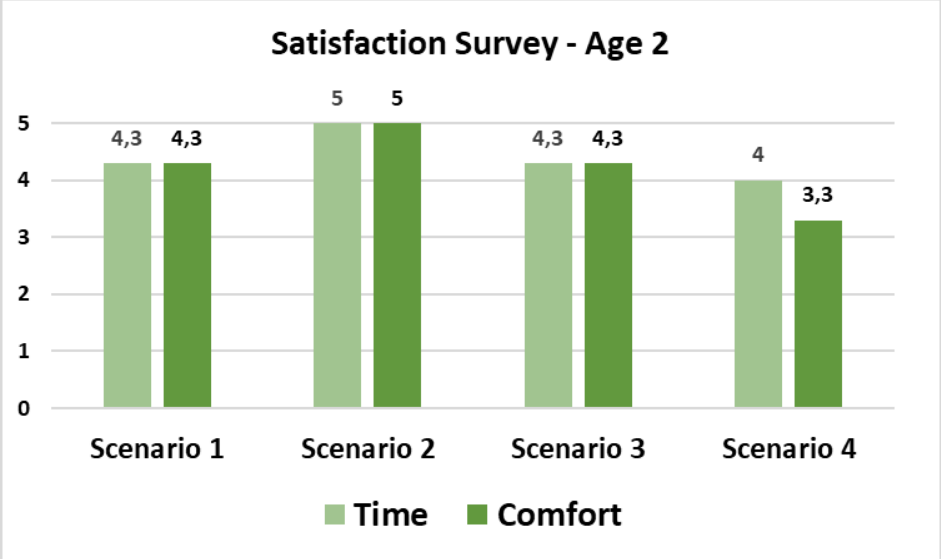


Figure 26: Satisfaction Survey results Age 2.

In the figure, we can observe the results of the satisfaction survey answered by the Age 3 group. Even in this case the scenario 2 obtained the highest score in terms of comfort and time. Additionally, 100% of them declared their general preference for the fingerprint recognition on mobile devices. Nevertheless, just 50% of this group stated their preference for using biometric authentication mechanisms in daily scenarios.

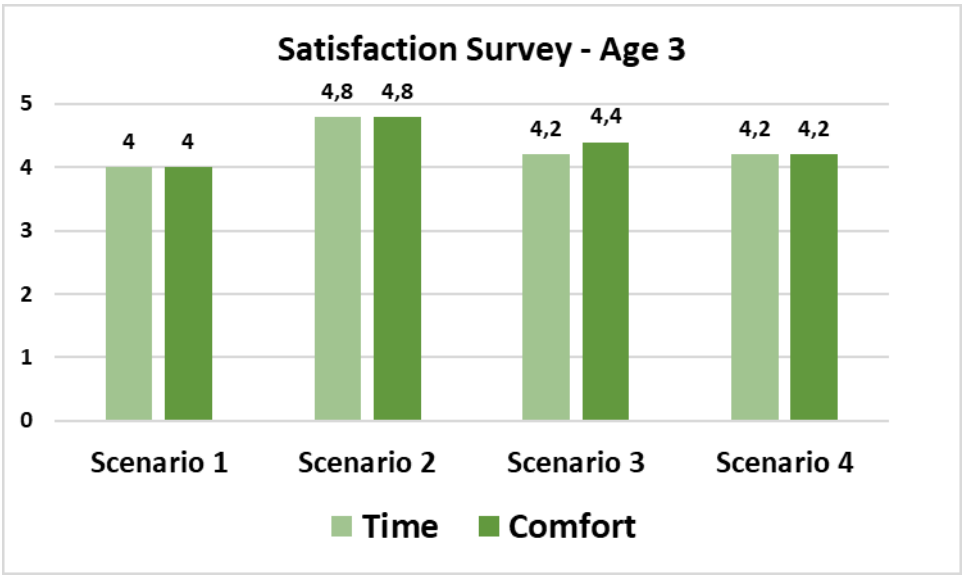


Figure 27: Satisfaction Survey results Age 2.

Finally, the satisfaction survey completed by the older users (Age 4 group) shows lower score rates for all scenarios (figure). Even if scenario 2 was rated as the faster scenario it was also the less comfort. Generally, elderly participants were not feeling at ease in completing the task required them alongside the evaluation. This is also confirmed by their preference. When we asked participants about their willingness of using biometric in access control scenarios, 75% declared that they will prefer using a traditional method such as keys.

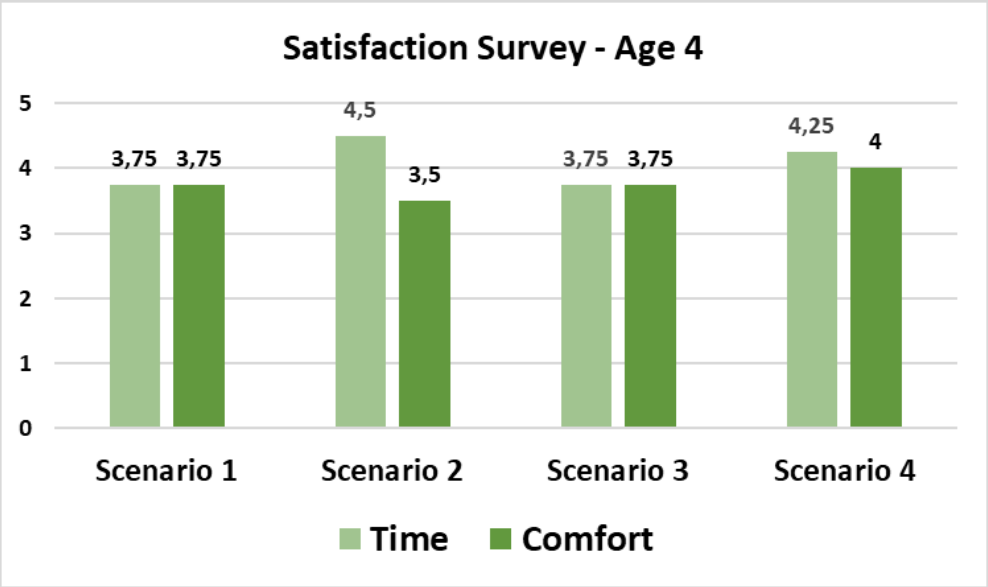


Figure 28: Satisfaction Survey results Age 4.

Regarding the satisfaction answers provided by the accessibility subgroups, the results are quite different compared with the control groups. Firstly, the comfort and the time needed for the four scenarios were rates by the accessibility groups with lower score rates compared with the scores given by control groups.

Users affected by developmental issues almost gave the same score rate to each scenario. Later, they expressed their preference for the IP camera (maybe because it is the scenario where is required less interaction between the user and the system). In addition, no users belonging to this group would be willing to use biometric to access their homes.

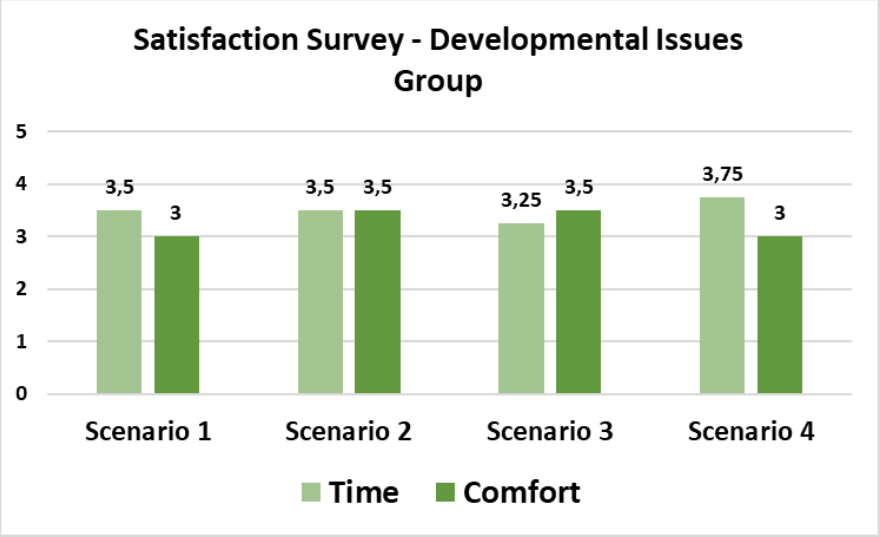


Figure 29: Satisfaction Survey results Developmental.

Participants, who took part in the experiment along with the learning issues subgroup, rated with low score the comfort of the second and four scenarios (Figure 29). This demonstrated that the learning issues impede users to feel at ease while completing mobile biometric recognition processes. Hence, the whole learning group stated to prefer the IP camera in access control scenarios. Even if, they did not know if they could be willing to use biometric applications in other scenarios.

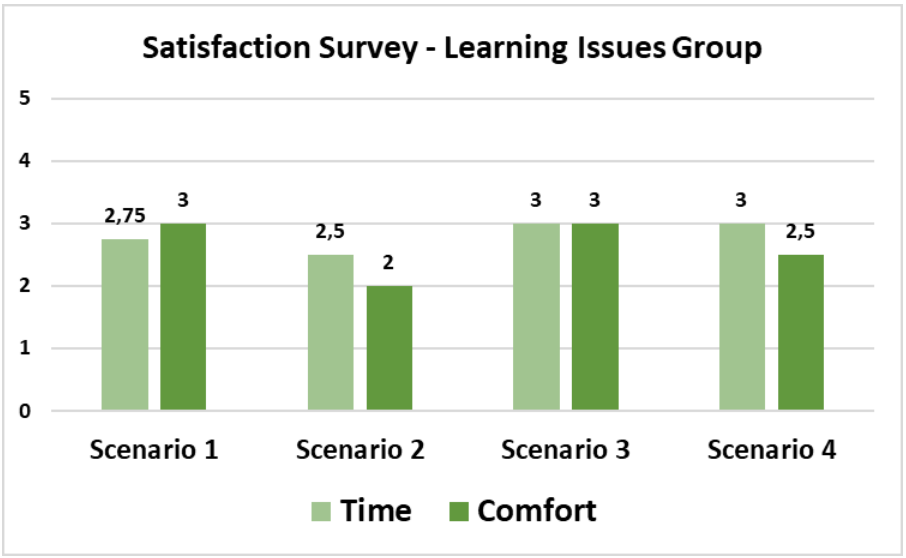


Figure 30: Satisfaction Survey results Learning Issues Group.

The score rates are higher when the motor issues subgroup answered the satisfaction questionnaire (Figure 31). Despite their motor concerns, 50% of these participants stated that they will be willing to use mobile biometric traits to complete daily tasks. While the other 50% were not sure to be willing to use biometric or mobile biometric tools.

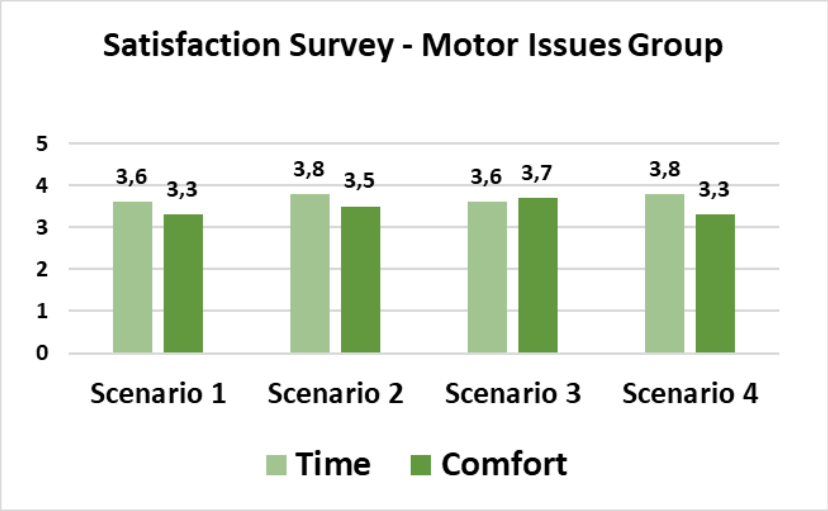


Figure 31: Satisfaction Survey results from Motor Issues Group.

6.1.4.3 Sample Quality

This last subsection focuses on the quality evaluation of the biometric sample captured during the accessibility evaluation. Thus, the fingerprint images (stored in Scenario 1) and the facial samples (captured during Scenario 3 and 4) were analysed according to the recommendation provided by the ISO/IEC TR 29794 part 4 and part 5.

The results are reported in the next subsections according to the biometric trait: fingerprint and face.

6.1.4.3.1 Fingerprint Sample Quality Analysis

During the first scenario, participants interacted with a capacitive fingerprint sensor (EikonTouch710 by Upek). As already explained, thanks to a specific C# application it was possible to extract the images and store them.

For each fingerprint sample, we evaluate the NFIQ1 score (since the samples were captured using a capacitive sensor it was not possible to report the NFIQ2 that works just with sample captured by an optical sensor) as a quality metric. Thus, this analysis was conducted using the NIST Fingerprint Minutiae Viewer [74]. This opensource software can return the NFIQ number for each fingerprint image uploaded in it. As already mentioned, the NFIQ1 values could be between 1 and 5; 1 is for high-quality fingerprint samples, while 5 for low-quality samples.

Figure 32 shows the NFIQ scores' distribution for the fingerprint samples collected by each population during the enrolment phase.

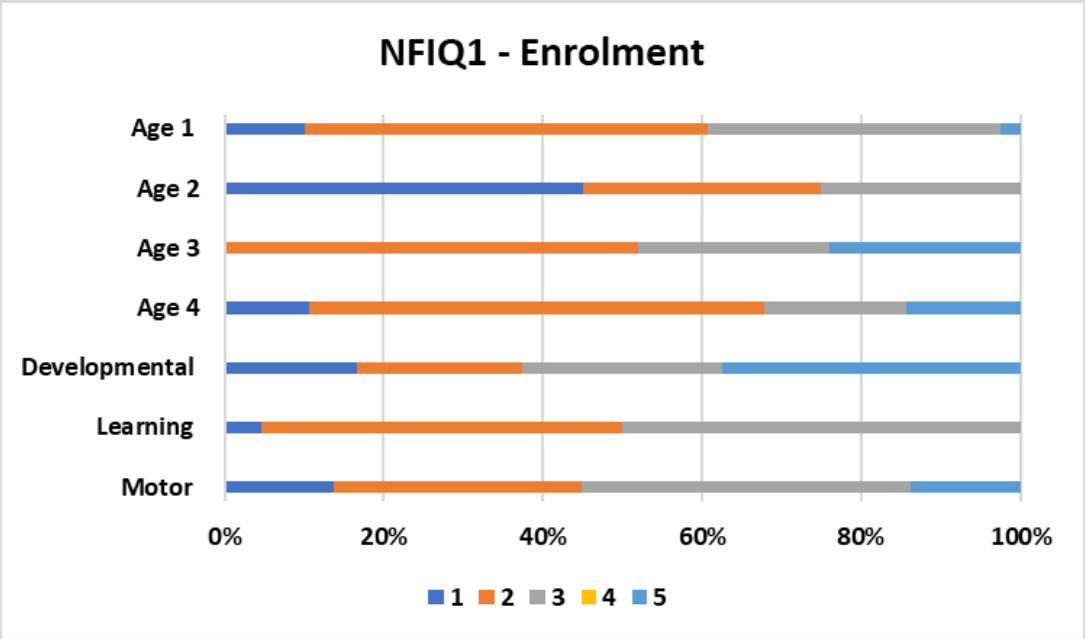


Figure 32: Distributions of NFIQ1 scores for the samples stored by each population during the enrolment.

As we can see from the figure above, the control populations mostly stored high-quality fingerprint samples. In fact, for this user sector, we have a high percentage of

fingerprint samples rated with the 1 and 2 values for the NFIQ1. Besides, it is also notable that users belonging to Age 3 and Age 4 stored more samples with NFIQ=5 compared with the younger subgroup. Thus, this confirms the correlation between age and the quality of the fingerprint images.

On the other hand, while the learning group collected samples with high-quality scores, this not happened for the image stored by the developmental and motor issues subgroups. The developmental group reached the highest percentage of samples with NFIQ1 = 5 in the enrolment phase. While the motor group presented the lower percentage of NFIQ1 = 1.

Regarding the first verification, the results are shown in figure 33.

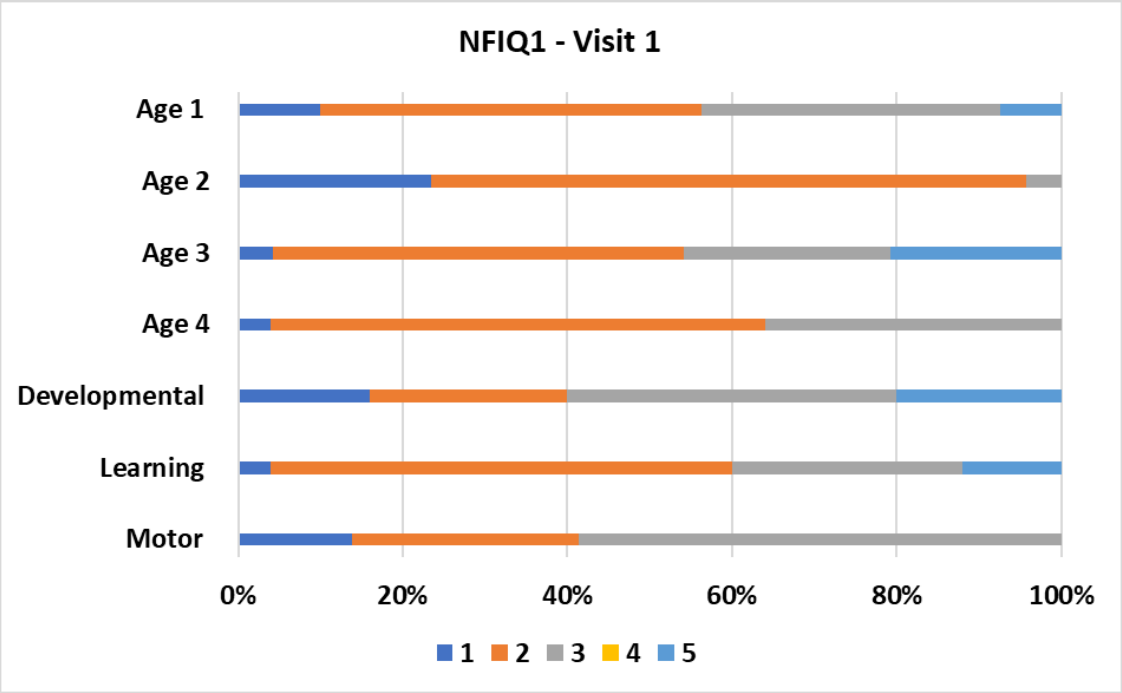


Figure 33: Distributions of NFIQ1 scores for the samples stored by each subgroup during the first visit.

The results are quite similar to the enrolment, except in some cases. For example, the percentage of samples with NFIQ1=5 decreased for the developmental group and it is null for the motor one.

Finally, in figure 34 we have the results of NFIQ1 scores for the second verification. The results are almost stable for the accessibility subgroups. While among the control users, we can notice that the Age 2 group provided a low percentage of NFIQ1 = 1 samples compared to the previous sessions. This could be addressed to the tiredness of the user at the end of the second visit.

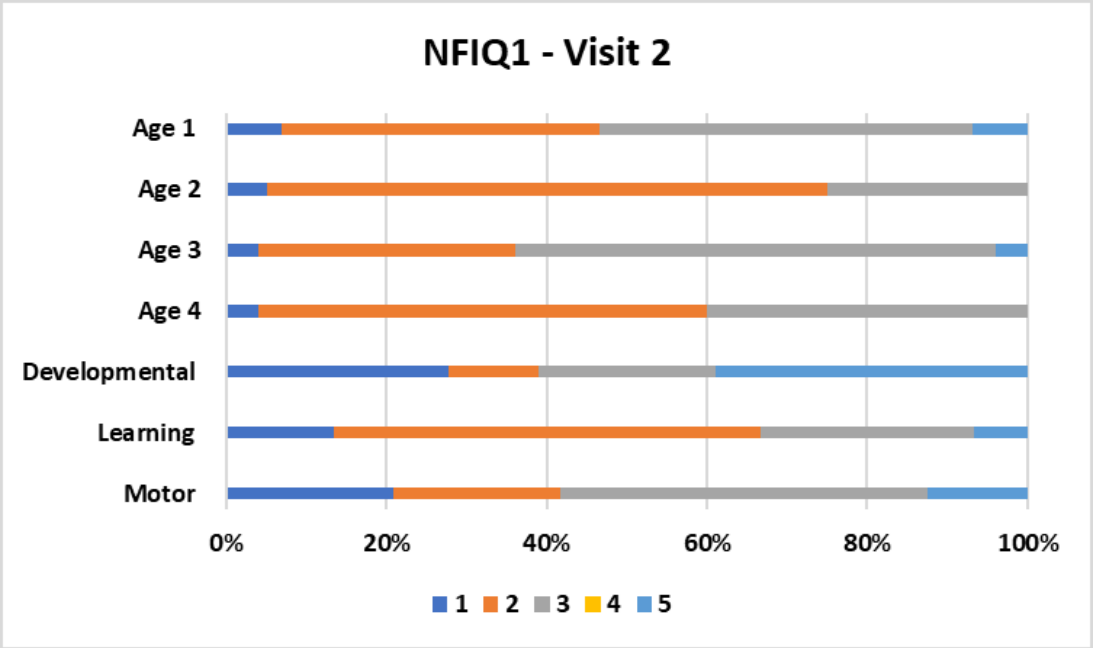


Figure 34: Distributions of NFIQ1 scores for the samples stored by each subgroup during the second visit.

6.1.4.3.2 Face Image Quality Analysis

The image captured with the IP Camera (Scenario 3) and by the Android app (Scenario 4) were also evaluated according to the quality analysis recommendations provided by the ISO/IEC TR 29794 – 5.

In particular, the assessment was conducted focusing on a specific directive enclosed in the ISO/IEC TR 29794-5. This standard suggests assessing the quality of the facial image evaluating the subject's behaviour. Thus, we assessed the face image quality by reporting the facial expression that each user had while completing the face recognition

sessions. We chose to focus on this specific recommendation to establish if age and accessibility concerns can even affect the users' feelings during in completing a face recognition process.

The face images captured in both scenarios (3 and 4) were processed again through VeriLook SDK, which is also able to detect the facial expression of the subjects.

A human being could show 7 different facial expressions: happiness, surprise, neutral, fear, disgust, anger, and sadness. These expressions are divided into positive (e.g., happiness, surprise, and neutral) and negative (such as fear, disgust, anger, and sadness) expressions.

While control users presented their face during the second scenario (Figure 35), the distributions of the facial expressions change among the different subgroups. Positive expressions are predominant in younger users (Age 1 and Age 2), while negative expressions characterized the face image collected with elderly users (Age 3 and Age 4).

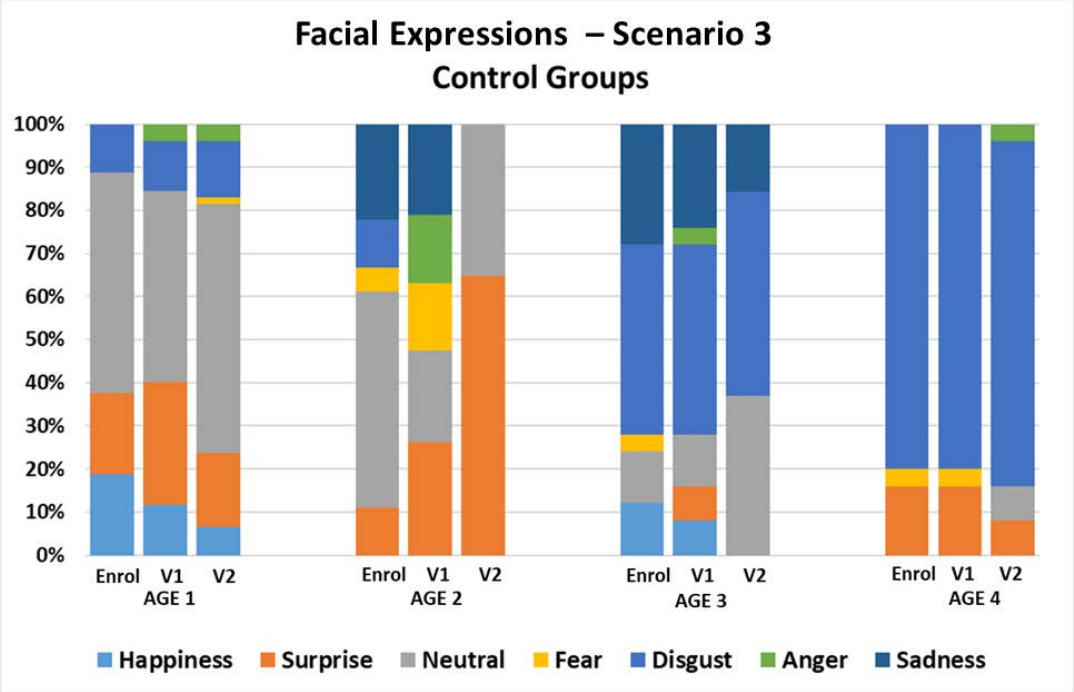


Figure 35: Distribution of Control Subjects' expressions while completing the tasks required in each part of scenario 2: enrolment (Enrol), first verification (V1), and second verification (V2).

Regarding the user affected by accessibility issues (Figure 36), positive expressions increased between the enrolment and the first verification meaning that the training lets users more willing to provide their facial traits. On the other hand, between the first and second visits, the distribution of positive expression is lower.

Another interesting result is that the most common expression was the surprise one. This could be addressed to the lack of experience of these subgroups in interacting with face recognition sensors. Hence, this subgroup did not know how to perform a face recognition process.

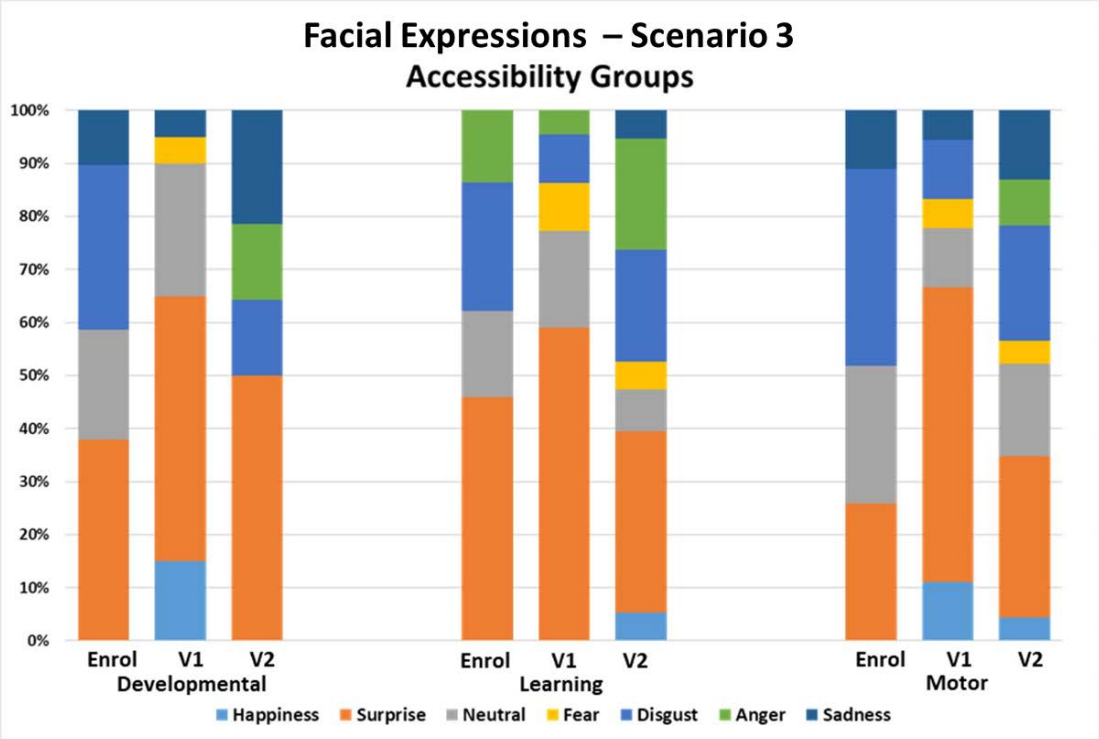


Figure 36: Distributions of Accessibility Subjects' expressions while completing the tasks required in each part of scenario 2: enrolment (Enrol), first verification (V1), and second verification (V2).

In the figure 37, there are the distributions of the control groups' facial expression performing the face recognition with the mobile application. Even in this case, younger users mostly took selfies with positive expressions. This is notable, especially in the Age 1 and Age 2 subgroups. The elderly group, instead, presented more negative expressions.

During the first and second verification, the users belonging to the Age 4 stored face image having just negative expression. In addition, in the older subgroups, the most predominant facial expression was the disgust face, meaning that this group was not at ease interacting with the face recognition on the smartphone.

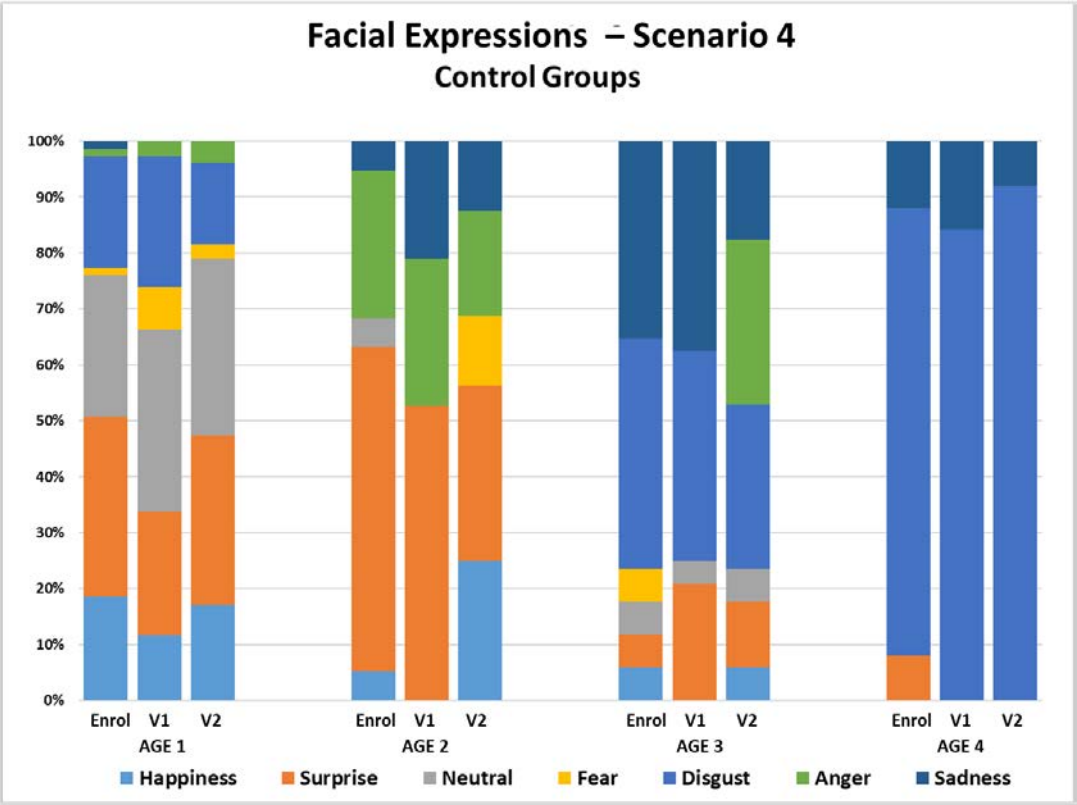


Figure 37: Distribution of Control Subjects' expressions while completing the tasks required in each part of scenario 4: enrolment (Enrol), first verification (V1), and second verification (V2).

Finally, users affected by developmental issues completed all the experiment's phases mainly with positive expressions (Figure 38). This did not happen with the learning issues subgroup which presented a higher distribution of negative expression when performed the first and second verification.

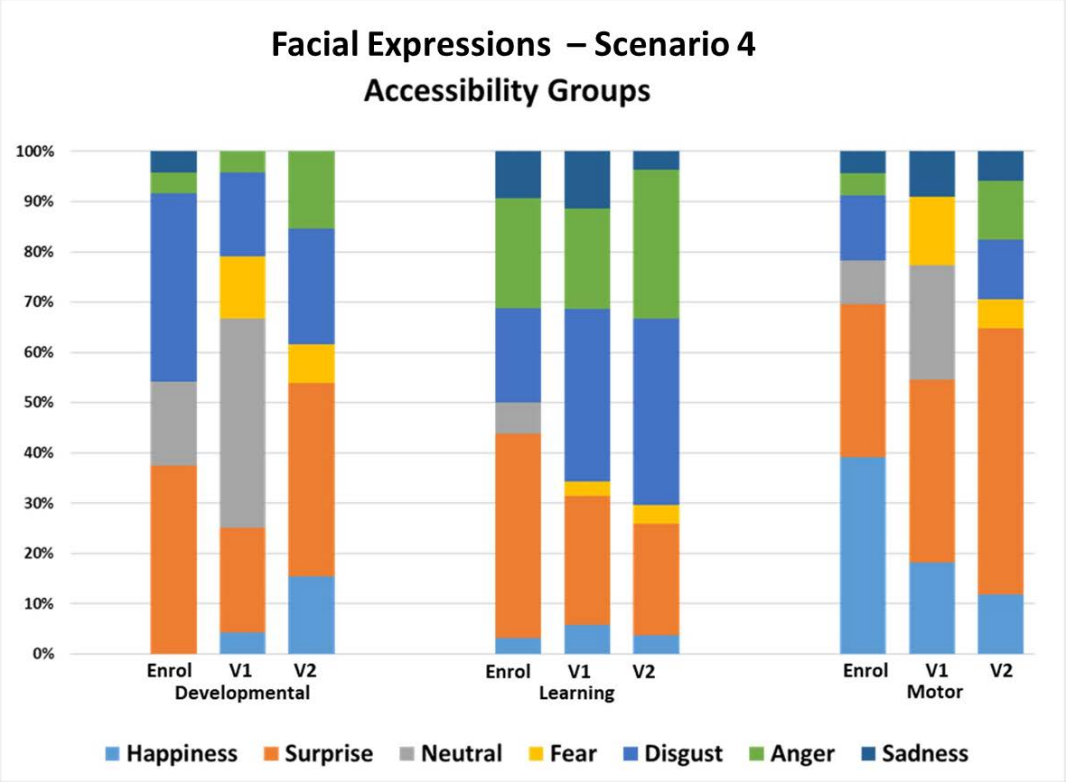


Figure 38: Distribution of Accessibility Subjects' expressions while completing the tasks required in each part of scenario 4: enrolment (Enrol), first verification (V1), and second verification (V2).

Additionally, Since the learning issues impeded the users to remember how to interact with the biometric application, the distribution of disgust and anger expressions increase between visit 1 and visit 2. While users with motor problems have mostly positive expressions while completing face recognition.

6.2 Overview of the results

This chapter reported the results for the accessibility test of an access control system based on fingerprint and face recognition. The analysis was conducted applying the novel methodology proposed in Chapter 4. Assessing each aspect this methodology suggests, it is possible to draw useful conclusions regarding the accessibility of the system.

Firstly, age affects the interaction between users and the biometric system., The number of negative expressions during the face recognition process with which user present their face to the IP camera (scenario 2) and the mobile camera (scenario 4) is noticeable. Elderly users also stated their preference to keep using traditional methods to access control systems instead of biometric recognition.

Developmental, learning, and motor issues also influence how the participants approached the biometric sensor during the scenario evaluation. Developmental issues impeded the users in understanding how to interact with the system. Users affected by learning issues have difficulties in the second visit in remembering how to present their biometric traits to the different sensors. While participants belonging to motor issues faced lots of drawbacks in interacting with the smartphone due to their low level of hand dexterity. Besides, looking at the satisfaction surveys, many users belonging to these subgroups stated that they do not know if they were willing to use biometric in real-life scenarios. This is an important aspect showing that there is a lack of knowledge regarding the application of biometrics and mobile in common scenarios.

Chapter 7 Mobile Fingerprint System for Retail Payments

This chapter details the results obtained when analysing the data collected during the second experiment. As already described, participants interacted with a fingerprint application that was running on three Android smartphones.

We required users to identify their fingerprint traits to complete a retail payment process, testing three scenarios evaluations. In each scenario, the data subjects interacted with a mobile device with the fingerprint sensor located on a specific side (front, lateral, and back) of the smartphone (Figure 39).

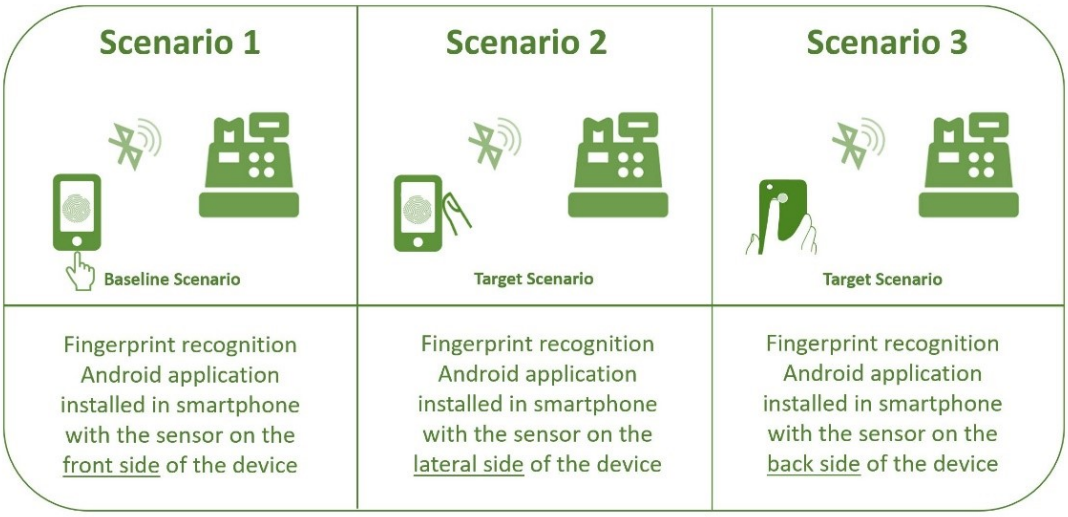


Figure 39: Scenario evaluation for testing the Mobile Fingerprint System.

Among the next sessions, the results are reported according to the novel methodology requirements (Chapter 4) observing each participant sector and each scenario separately.

Finally, the chapter ends with the discussion regarding the results obtained by the different user sectors to provide an overview regarding the accessibility of the system.

7.1 Result Analysis according to the Novel Methodology

As suggested by the novel methodology, the data collected during the experiment were analysed to report three different information:

- The accessibility of the scenario
- The accessibility of the biometrics system
- How the accessibility concerns impact on the outcome of the biometric process

These aspects will be discussed in the next subsection after providing a brief description of the enrolled data crew.

7.1.1 Data Crew

As stated in Chapter 5, for this experiment 21 users were recruited. The whole data crew was split into a control and an accessibility group.

The control group is comprised of 6 users without any kind of accessibility issues. According to the user age, this group was split into 2 subgroups: Age 1 and Age 2.

The first subgroup (Age 1) was composed of 4 users (1 female and 3 male), between 24 and 30 y/o. The second subgroup (Age 2) gathered 2 users (1 male and 1 female), both older than 45 y/o.

The data subject of the accessibility sector gathered 16 users affected by different types of accessibility issues. This group was divided into three smaller groups according to the participant accessibility concerns.

The characteristics of each group are summarized and shown in table 22:

Table 22: Characteristics of the users enrolled in the second evaluation.

Group	Subgroup	Characteristics	Number of Users
Control	Age 1	24 – 30 years old	4
	Age 2	45 + years old	2
Accessibility	Developmental	Congenital related disorders affecting the cognitive ability	5
	Learning	Intellectual disorders affecting by Attention Deficits	10
	Motor	Affected by hand concerns (1 user affected by hand arthrosis) and 2 of them by leg issues (1 participant was using crutches)	3

7.1.2 Accessibility of the Scenario

In this part, we are going to analyse the accessibility of the scenario. As proposed along with the novel methodology, this aspect is reported by observing whether the users can have access to the scenario evaluation and if they have already experienced the required task in real-life scenarios.

Regarding the control groups, all users from Age 1 and Age 2 were able to assess the scenario recreated for the accessibility evaluation.

The 4 users, belonging to Age 1 group, already had experience with mobile payment apps. Besides, three of them declared to be already using fingerprint recognition to unlock their smartphones. While no participants belonging to the Age 2 group had experience with any similar applications and they were not used to biometric processes on mobile devices.

Besides, volunteers belonging to developmental, learning, and motor issues group could assess the scenario recreated for the evaluation of the mobile fingerprint system. Some participants were also smartphone users (table 23). While none of them had previous experience with mobile biometric applications.

Table 23: Users' experience with smartphones and mobile biometrics.

Subgroup	Smartphone Owners	Mobile biometric experience
Age 1	4/4	4/4
Age 2	2/2	0/0
Developmental	2/5	0/5
Learning	5/10	0/10
Motor	2/3	0/3

7.1.3 Accessibility of the system

This section deals with the accessibility of the system. Thus, the following table (table 24) reports the number of users that cannot start interacting with the system in each scenario.

Table 24: Number of users who cannot interact with the system.

Subgroup	Scenario 1 Device 1 (D1)	Scenario 2 Device 2 (D2)	Scenario 3 Device 3 (D3)
Age 1	0	0	0
Age 2	0	0	0
Developmental	2	1	2
Learning	1	2	3
Motor	0	1	1

Users, coming from Age 1 subgroup, were able to interact with the system during each part of the scenario evaluations. Regarding the Age 2 subgroup, one participant could not enrol his thumb finger on any smartphone. This was due to the poor quality of the finger skin, confirming that age affects the quality of the fingerprints.

Developmental issues impeded 2 users to interact with the D1 and D3 and 1 user with the D2.

While participants belonging to the learning issues group had several difficulties in understanding, firstly, where the fingerprint sensors were and, secondly, how to present the fingerprint to store the templates. Thus, one user did not start the interaction with D1, and two users cannot interact with D2 and D3.

Due to hand arthrosis, a user belonging to the motor-concerns group could not start the interaction with the D2 and D3.

Table 25 reports the number of users who could not complete any part of the evaluation.

Table 25: Number of users who cannot complete the task required in the scenarios.

Subgroup	Scenario 1 Device 1 (D1)	Scenario 2 Device 2 (D2)	Scenario 3 Device 3 (D3)
Age 1	0	0	0
Age 2	0	0	0
Developmental	2	1	2
Learning	1	2	0
Motor	0	1	1

All users belonging to Age 1 and Age 2 subgroups complete the task required them during each part of the evaluation.

Developmental and learning issues caused a lack of memory in users. Hence, some participants belonging to these two subgroups cannot complete the interaction because they did not remember how to interact with smartphones.

Besides, one user coming from the motor group was not able to complete the interaction with D2 and D2 due to a lack of hand dexterity.

7.1.4 How the accessibility concerns impact on the outcome of the biometric process

This part discussed the third step of the methodology presented in this Thesis. Thus, the result discussion presented alongside the following parts aims to understand if there is any correlation between the age and health status of the users and their interaction with the biometric recognition system.

According to the methodology proposed in Chapter 4, we are going to report the performance and usability scores recorded during the evaluation. For this experiment, the quality of the fingerprint samples cannot be reported since the Android fingerprint security settings do not allow image extraction.

7.1.4.1 Performance

The performance of the system was analysed reporting the percentage of the successful recognition attempts done by the users during the two verification visits.

During the first visit, the Age 1 group obtained the best performance score interacting with the second smartphone (D2). While the interaction with the third device (D3) brought the lowest performance scores (Figure 40). This can be explained considering the user's tiredness at the end of the first visit.

Between the sessions, the performance scores increased just for the first devices (D1). As illustrated in the figure, during the second visit, all the D1 verification attempts had a successful outcome.

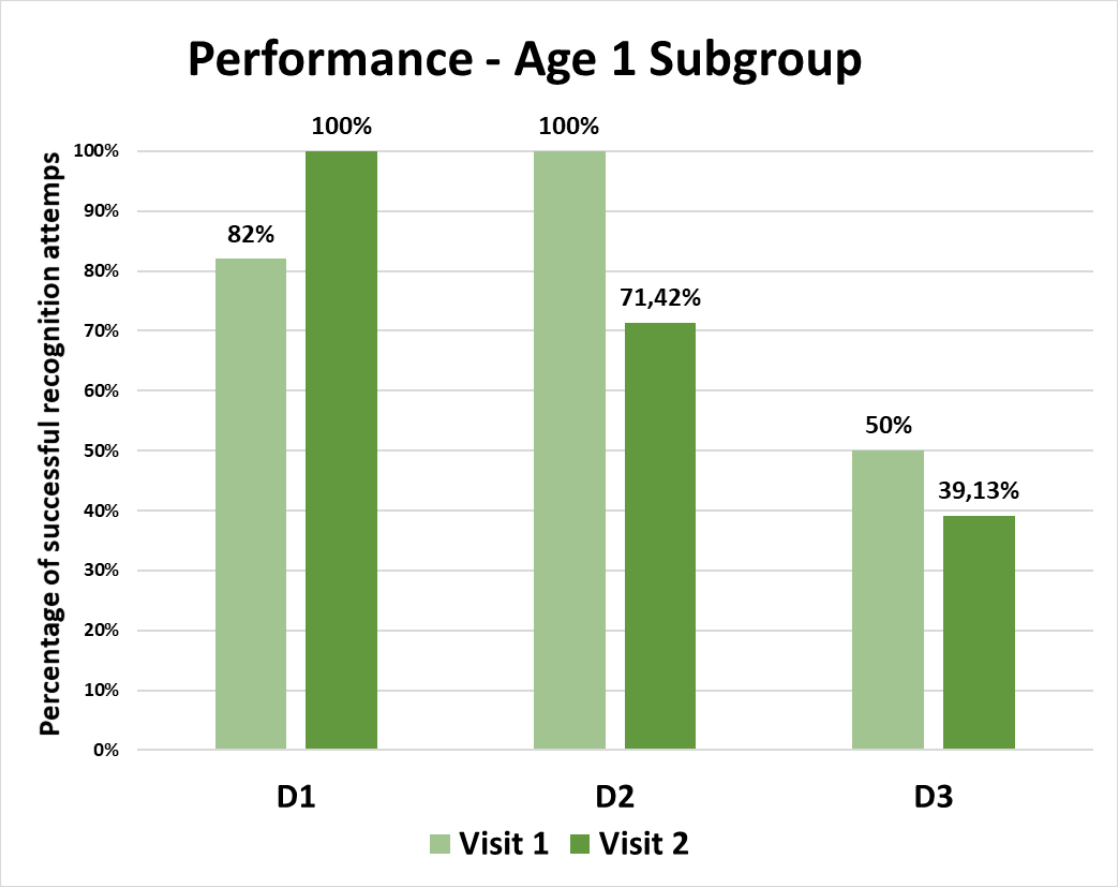


Figure 40: Percentage of successful recognition attempts obtained by Age 1 group interacting with D1, D2, and D3 during the first and the second visit.

It can be observed that the Age 2 group obtained different performance scores compared to the first group (Figure 41). This time, the best device (in terms of performance) was the D1 in both visits. Besides, D2 and D3 had lower performance outcomes on both visits. This can be addressed to the lack of hand dexterity of the users while interacting with the sensor in the lateral and backside of the smartphones.

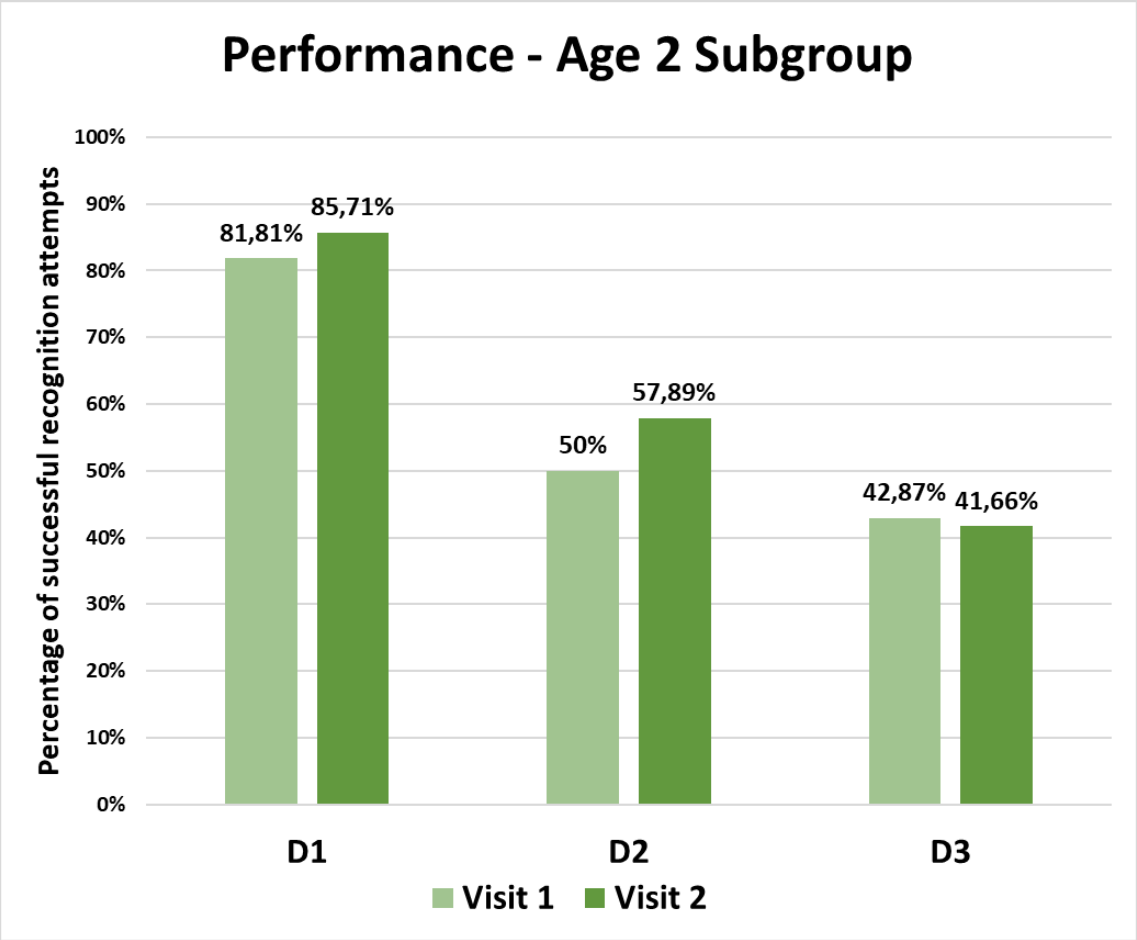


Figure 41: Percentage of successful recognition attempts obtained by Age 2 group interacting with D1, D2, and D3 during the first and the second visit.

Comparing the results obtained from both groups, it is noticeable that the Age 2 did not reach results as high as Age 1 group. Thus, this can be attributed to the quality of the biometric samples, to the poor experience and dexterity of the elderly users.

Regarding the accessibility group, users affected by developmental issues gained experience between the sections. The percentage of successful recognition attempts increases from visit 1 and visit 2 (Figure 42) during the first and the third scenario. Regarding the second scenario, the percentage of successful attempts remained almost constant during the two visits.

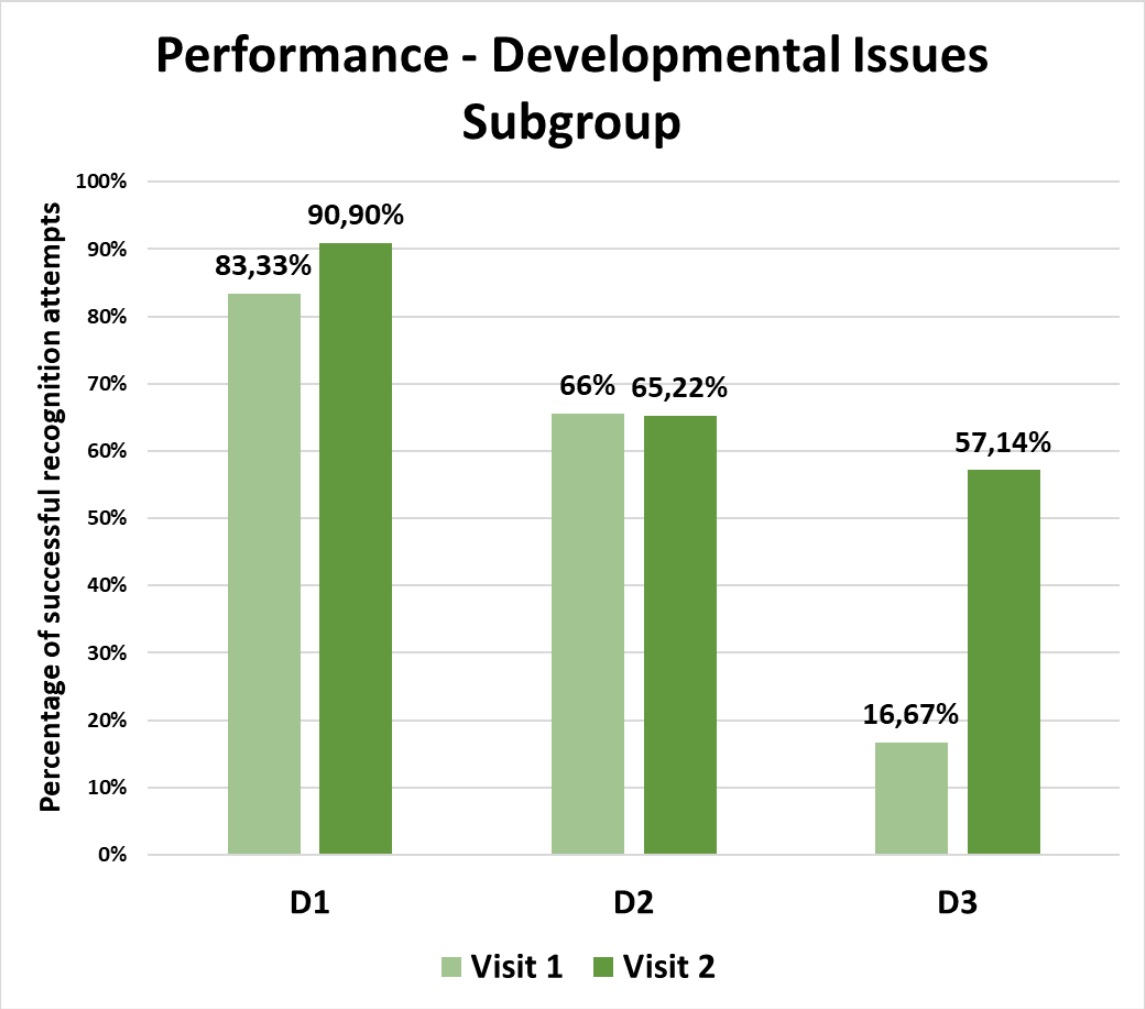


Figure 42: Percentage of successful recognition attempts obtained by developmental issues group interacting with D1, D2, and D3 during the first and the second visit.

The worst performance scores were recorded while users tested the third scenario, and it was due to the configuration of the fingerprint sensor of D3. Participants had several

interaction issues in presenting the fingerprint without being able to see the biometric sensor (located in the backside of the smartphone).

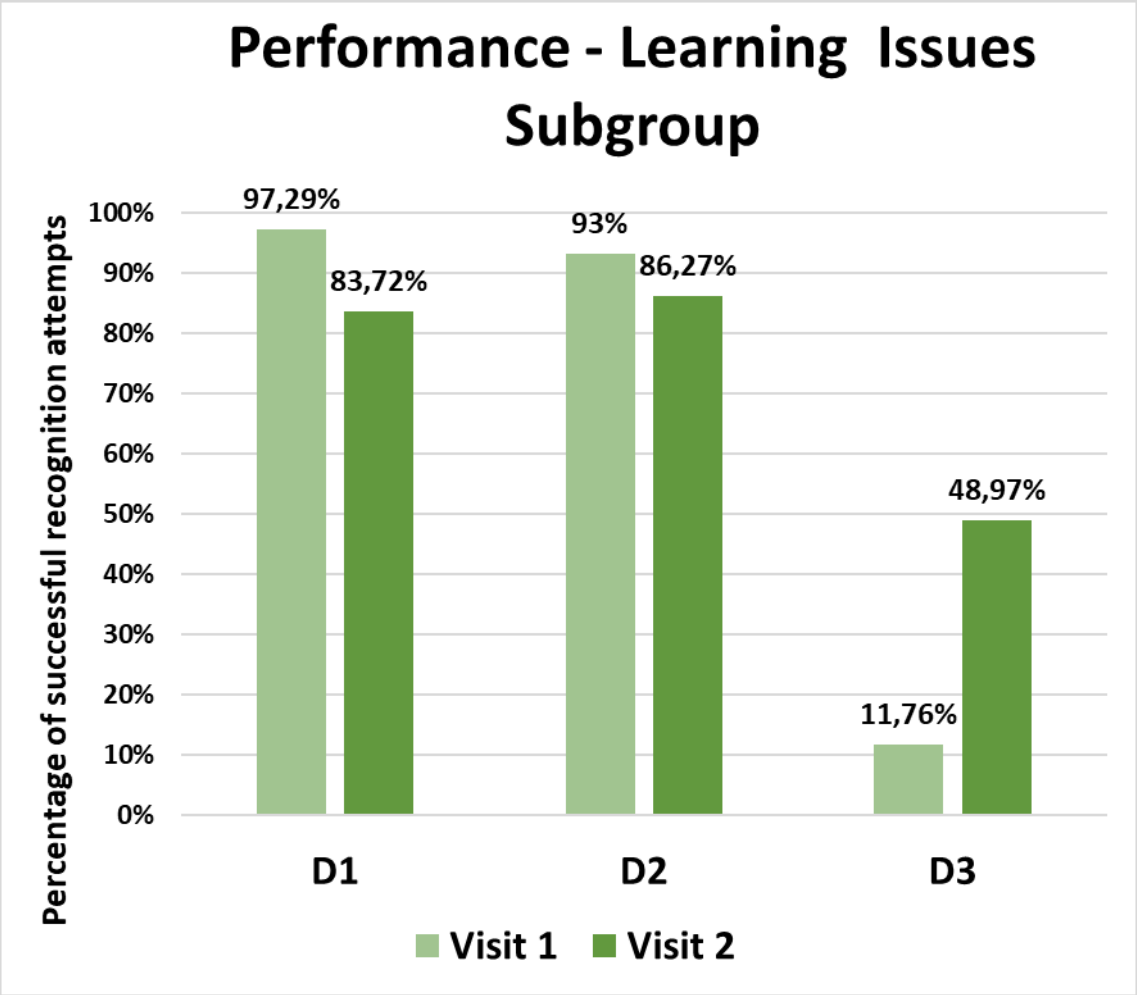


Figure 43: Percentage of successful recognition attempts obtained by learning issues group interacting with D1, D2, and D3 during the first and the second visit.

Participants with learning issues registered a decrease of performance scores among first and second visit with D1 and D2 (Figure 43). This can be addressed to the difficulty of the volunteers to remember where the sensor was positioned and how to interact with it.

A low percentage of successful recognition attempts were reached in the first verification of D3 may be due to the tiredness of the user at the end of the first verification session. In fact, the performance of the third scenario increased in the second visit when the users interacted with the D3 at the begging of the visit.

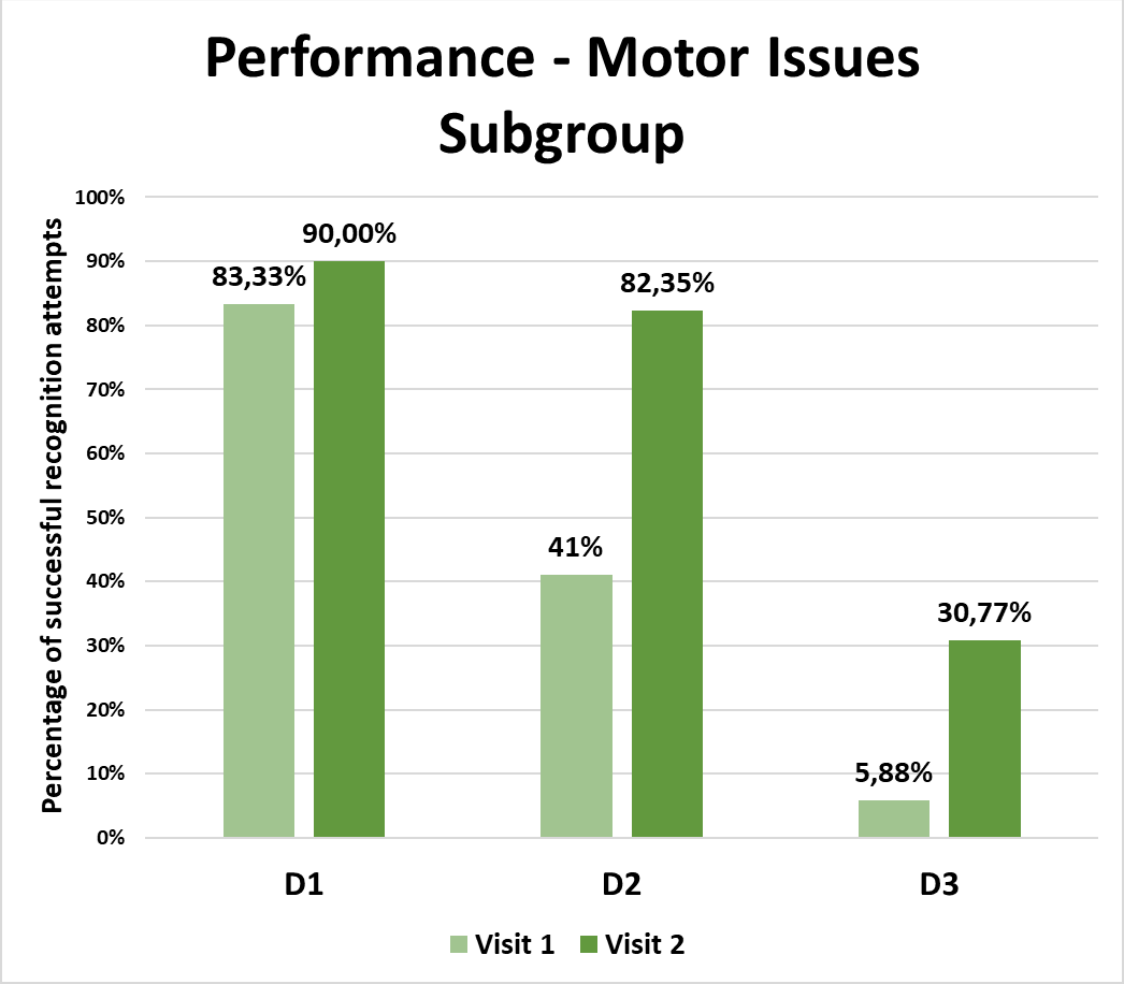


Figure 44: Percentage of successful recognition attempts obtained by motor issues group interacting with D1, D2, and D3 during the first and the second visit.

Due to their motor problem, the third group obtained a low level of percentage interacting with the D3 (Figure 44). For this group, it was quite difficult holding the device and touching the backside fingerprint sensor. The percentage of genuine attempts increased between the session reaching the value of 30,77% (the lowest level of performance in the second visit).

The interaction with the D1 and the D2 brought better results compared with the third scenario. D1 was the best device in terms of performance scores. The percentage of genuine attempts in the first and second scenarios increased alongside the second visit especially interacting with the D2.

7.1.4.2 Usability

The usability is going to be analysed according to the metrics specified in Chapter 4, this means reporting the efficiency, effectiveness, and satisfaction.

7.1.4.2.1 Efficiency

The efficiency was measured reporting the seconds spent by the user during the enrolments and the verifications.

When considering the efficiency in the first enrolment of the evaluation, the results are very changing depending on the user subgroups (Table 26).

Table 26: Mean (μ) and standard deviation (σ) of the time (in second) spent by Age 1 during the first enrolment.

Subgroup	Device 1 (D1)				Device 2 (D2)				Device 3 (D3)			
	Index		Thumb		Index		Thumb		Index		Thumb	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
Age 1	15,97	5,47	19,36	4,67	18,59	14,69	29,94	27,72	22,95	14,36	24,63	17,21
Age 2	44,45	26,1	62,8	-	55,18	23,4	45,5	-	18,97	7,03	33,5	-
Developmental	36,58	36,5	20,62	44	22,38	7,4	34,33	30,3	31,29	19,14	23,72	7,07
Learning	32,82	16,7	28,89	21,09	29,05	8,87	19,2	3,01	36,18	7,9	17,07	2,9
Motor	25,11	0,6	22,64	1,76	27,44	5,7	37,3	27,61	14,36	-	21,25	10,54

During the first enrolment, participants belonging to the Age 1 group took more time in storing their fingerprint traits on the D2 and D3 compared to the first scenario (interacting with the D1). This depended on the position of the sensor (lateral and back) that sometimes confused the users. The developmental group spent more time in completing the enrolment with the D1, while learning and motor group with the D3 and D2 respectively.

When in the second visit Age 1 group enrolled their fingerprint with the D2 and D3, they spent less time compared with the efficiency scores of the first visit (Table 27). The second enrolment of D1 took more time compared with the first one. This probably depended on the users' tiredness at the end of the second verification.

Table 27: Mean (μ) and standard deviation (σ) of the time (in second) spent by users during the second enrolment.

Subgroup	Device 1 (D1)				Device 2 (D2)				Device 3 (D3)			
	Index		Thumb		Index		Thumb		Index		Thumb	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
Age 1	24,6	7,68	35,4	15	9,7	5,65	12,4	6,45	18,54	4,82	12,7	2,4
Age 2	31,87	13,54	75	-	44,88	23,22	41,2	-	49,03	-	32,9	-
Developmental	36,67	18,23	27,7	17	34,98	27,8	35	18,48	17,41	6,88	35	18,48
Learning	33,89	19,55	27,2	10,2	27,5	4,6	24,08	5,73	20,19	5,44	21,62	5,94
Motor	26,68	5,4	23,6	0,69	29,7	3,35	22,9	2,79	23,41	1,6	23,9	3,46

In all scenarios, the Age 2 group took more time to enrol their fingerprint templates compared with the first group. This happened also in the second session of the experiment (second enrolment phase) when the users were asked to provide again their fingerprint.

In the second enrolment, the interaction time decreased for all the users just during the third scenario. All participants took more time to enrol their fingerprints alongside scenarios 1 and 2. This is probably due to the order in which the user completed the evaluations.

During the second session, volunteers started the experiment with the third device, thus, they were less tired in interacting with the D3 smartphone.

In addition, looking at tables 26 and 27, enrolling the thumb generally took more time than the index finger.

Regarding the efficiency of the verification phases (table 28), users from Age 1 group completed the second scenarios 1 and 2 faster than the first visit. Thus, experience helps users in interacting with mobile biometrics, at least for younger users. Compared with the first visit, Age 2 group spent more time completing the three scenarios in the second visit. This is probably because the users tended to forget the position of the sensor in each scenario.

Table 28: Mean (μ) and standard deviation (σ) of the time (in second) spent users during the verifications.

Subgroup	Device 1 (D1)				Device 2 (D2)				Device 3 (D3)			
	Visit 1		Visit 2		Visit 1		Visit 2		Visit 1		Visit 2	
	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ	μ	σ
Age 1	8,8	5,1	5,3	2,9	8,6	4,4	9,8	5,3	16,4	9,9	11,8	6,2
Age 2	4,5	2,1	6,2	7,8	4,9	1,5	7,2	2,9	10,1	13,9	8,5	-
Developmental	9,9	2,9	8,9	3,5	10,8	3	7,1	1,5	9,8	4,2	7,6	1,4
Learning	8,5	2,3	9,4	3,7	8,2	2,1	8,8	3,3	8,5	5	10,3	2,8
Motor	9	1,6	6,3	1	7,6	-	7,3	1,6	12,6	7,6	7,6	1,3

When assessing the verifications completed by the accessibility groups, for the developmental and motor group the interaction time decreased from the first verification to the second one. While the learning issues group spent more time during the second visit than in the first one. This depended on the uncertainty of the users to complete the second visit due to not remembering how to interact with the smartphones.

While assessing the efficiency with which each group completed the various stages of the test, we conduct the ANOVA test to verify the statistical independence between the 5 subgroups analysed. Through the ANOVA test, we obtained: 0,0037 (Scenario 1), 0,0021 (Scenario 2), 0,0086 (Scenario 3), and 0,000269 (Scenario 1), 0,003 (Scenario 2), 0,05 (Scenario 3) for the first enrolment of the index and the thumb respectively; 0,019 (Scenario 1), 0,05 (Scenario 2), 0,032 (Scenario 3), and 0,019 (Scenario 1), 0,00326

(Scenario 2), 0,0369 (Scenario 3) for the first enrolment of the index and the thumb respectively. While analysing the efficiency of the verifications the p-value scores are: 0,032 (Scenario 1), 0,05 (Scenario 2), 0,02 (Scenario 3) for the visit 1; and, 0,00096 (Scenario 1), 0,0223 (Scenario 2), 0,0323 (Scenario 3) for the visit 2. Thus, the p-values obtained are lower or equal to the null hypothesis value (0,005) confirming that subgroups are statistically independent.

7.1.4.2.2 Effectiveness

The effectiveness was evaluated through the percentage of incorrect interactions made by the user in the first and second verification (Table 29).

Table 29: Percentage of incorrect interactions made by each group during the verifications.

	Device 1 (D1)		**Device 2 (D2)**		**Device 3 (D3)**	
Subgroup	**Visit 1**	**Visit 2**	**Visit 1**	**Visit 2**	**Visit 1**	**Visit 2**
Age 1	13,04%	5,88%	0%	5%	4,16%	11,11%
Age 2	18,18%	21,42%	30,76%	15,78%	85,71%	50%
Developmental	16,21%	2,32%	4,4%	19,6%	17,64%	8,16%
Learning	11,1%	9,09%	6,89%	8,69%	36,6%	14,28%
Motor	16,6%	10%	5,12%	35,29%	35,29%	15,38%

Younger users did fewer interaction errors when completing the second scenario, and they generally made fewer mistakes in both visits compared with older people. Older participants made a lot of incorrect interaction especially in the third scenario: during the first visit, they made 85,71% of incorrect interaction. Even if during the second visit the percentage of incorrect interactions decreased meaning that the user got experience in interacting with the back-side fingerprint sensors.

Between the visits, developmental and learning issues groups gained experience on how to present the fingerprint to the mobile biometric sensors. In fact, during the second verification, both groups made less incorrect interactions. Participants with motor issues also gained experience (when interacting with the fingerprint sensor) during the sessions interacting with the D1 and D3. While the interaction with the D3 in the second visit was quite difficult for this group which made 35,29% incorrect interactions.

7.1.4.2.2 Satisfaction

The satisfaction was assessed asking participants several questions regarding the scenarios and the interaction with the application using the three different smartphones.

Firstly, users were required to complete a satisfaction survey to rate the comfort, the interaction time, and the easiness of each device declaring which smartphone they preferred during the completion of the experiment. Secondly, they were asked regarding their wiliness of using mobile biometrics in a real-life context.

As concerned with satisfaction survey Age 1 group declared the D1 as the best device in terms of comfort, time, and easiness; followed by the D3 and D2 (Figure 45).

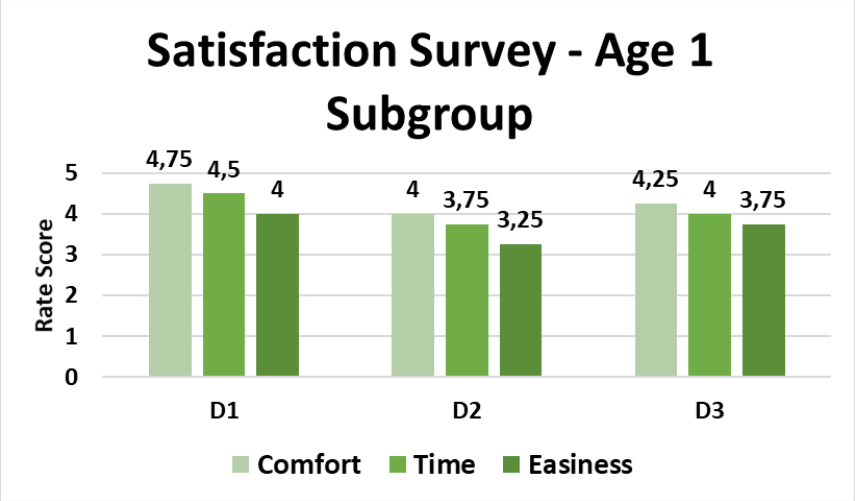


Figure 45: Satisfaction Survey results of the Age 1 group.

D1 was the best device even for Age 2 (even if the older group rated the D1 with lower scores compared with the younger users). D2 and D3 were rated with the same level of scores (Figure 46).

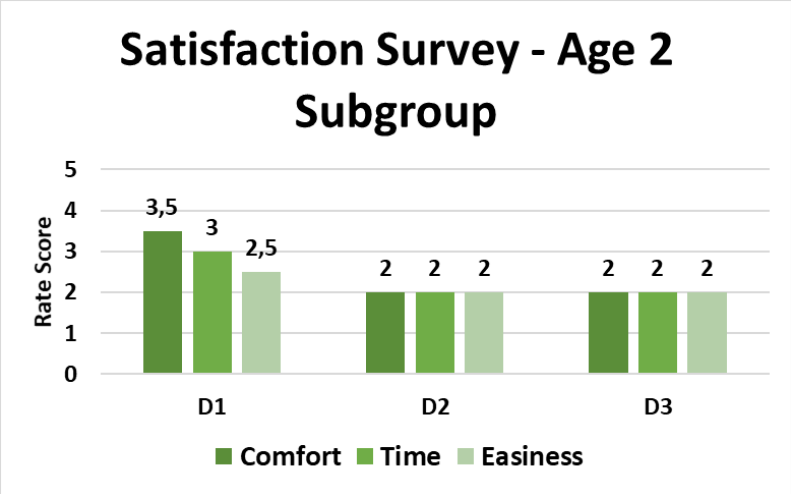


Figure 46: Satisfaction Survey results of Age 2 group.

People with developmental concerns rated the D1 as the best device in terms of time and the D2 as the most comfortable and easy-to-use (Figure 47). While the D3 obtained lower scores of the satisfaction respect the other 2 devices.

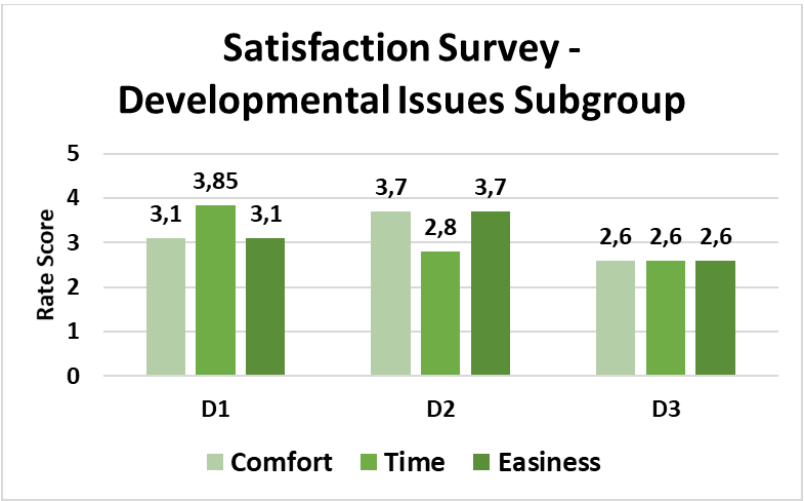


Figure 47: Satisfaction Survey results of the Developmental Issues subgroup.

The D1 was the best device in terms of comfort and time for the users with learning issues (Figure 48). Besides, the D3 was rated with the highest scores as regarding the easy-to-use.

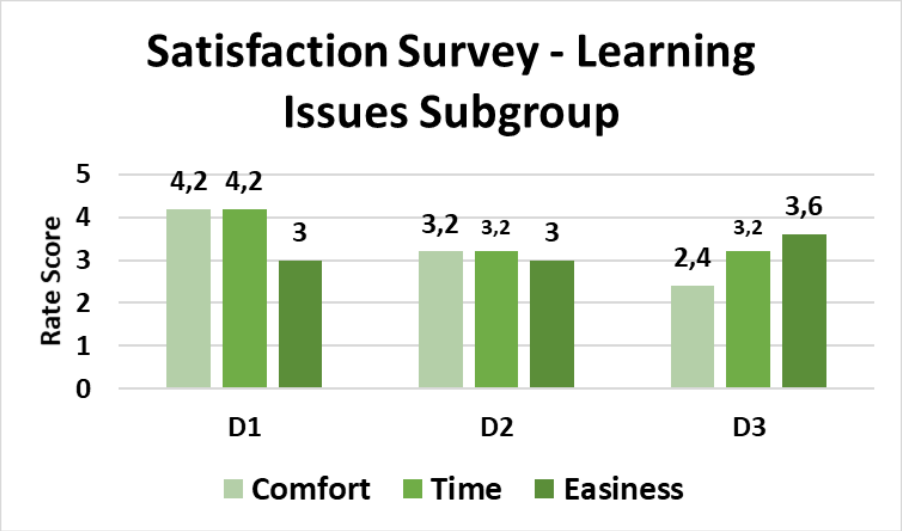


Figure 48: Satisfaction Survey results of the Learning Issues subgroup.

Participants with motor issues rated the D1 as the best in terms of comfort, time, and easiness (Figure 49). The D2 and D3 were rated with low scores, meaning that users found very difficult interacting with these two devices.

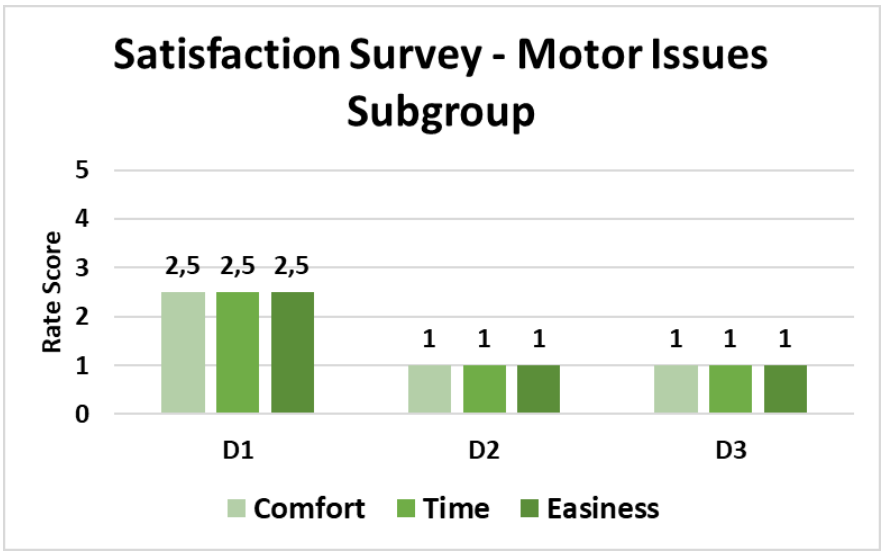


Figure 49: Satisfaction Survey results of the Motor Issues subgroup.

Regarding the favourite smartphone among the three devices user during the evaluations.

75% of Age 1 volunteers chose the D3 and 25% the D2 (Figure 12.a). While the two users who made up the Age 2 group, both declared the D1 as their favourite smartphone (Figure 50.b).

Figure 50: Favourite Device for Control groups. a) Age 1 group favourite mobile device b) Age 2 group favourite mobile device.

The first device was also rated as the favourite device by most of the users from all the accessibility groups. 71,42% of participants, belonging to the developmental issues group, chosen the D1, while 14,28% the D2, and another 14,28% the D3 (Figure 13.a).

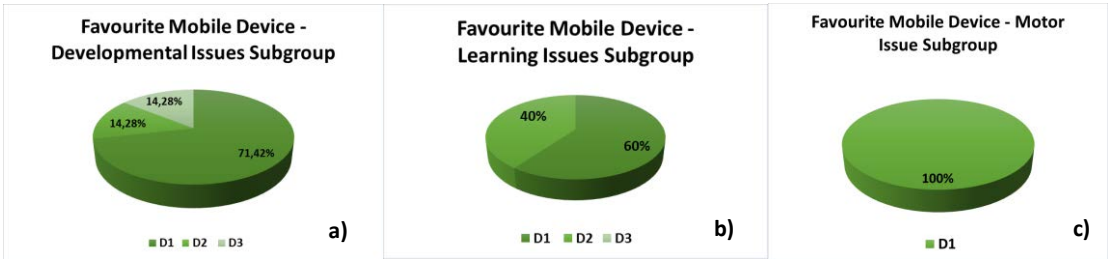


Figure 51: Favourite Device for Accessibility groups. a) Developmental Issues favourite mobile device, b) Learning Issues favourite mobile device, b) Motor Issues favourite mobile device.

60% of the second accessibility group rated the D1 as the favourite smartphone, the remaining 40% preferred the D2 (Figure 13.b). While, the participants with motor issues declared the D1 as their favourite smartphones (Figure 13.c).

We also asked if they would have been willing to use the system tested in real retail payment scenarios. The 75% answered yes (Figure 14.a), instead of the 25% who would prefer to use other payment methods. While just 50% of Age 2 participants would be willing to use mobile biometrics in retail payment contexts (Figure 14.b).

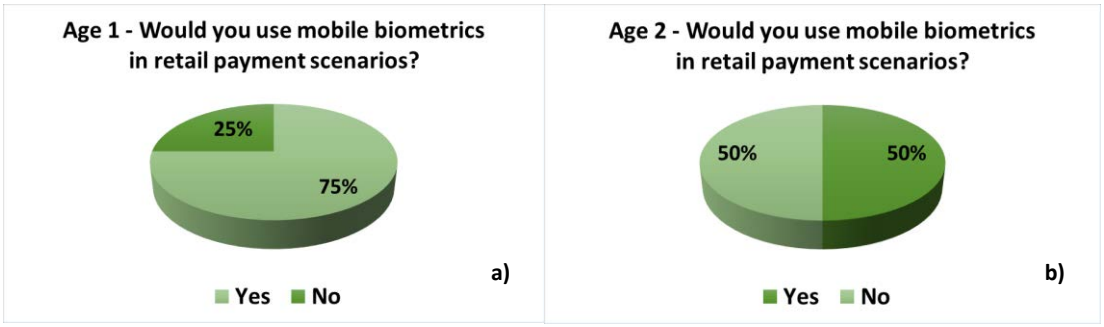


Figure 52: User opinion regarding the application of mobile biometric in real retail payment scenarios a) Age 1 group answers, b) Age2 answers.

Finally, most of the users belonging to the developmental and learning issues groups stated that they would like to use the mobile biometric application in retail scenarios (Figure 15.a and 15.b).

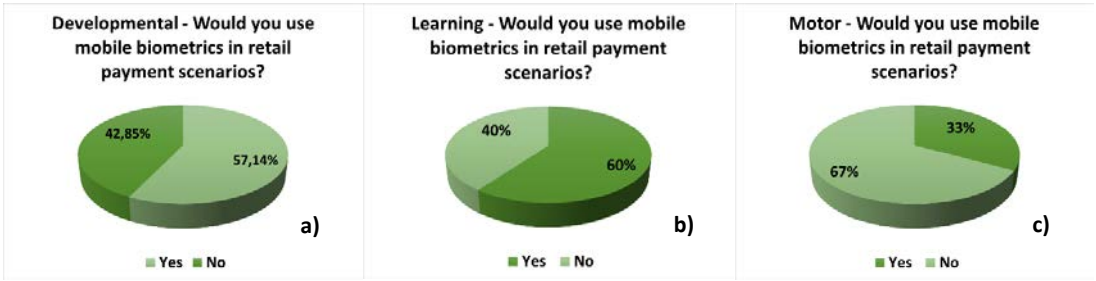


Figure 53: User opinion regarding the application of mobile biometric in real retail payment scenarios a) Developmental Issues group answers, b) Learning Issues group answers, c) Motor Issues group answers.

Just 33% of the motor issues group (one user out of three) said to prefer paying with mobile biometric instead of the traditional payment methods (Figure 15.c).

7.2 Overview of the results

This chapter provided the results obtained analysing the accessibility of a mobile biometric application for retail payments. The accessibility of this system was reported according to the formal methodology proposed in chapter 4.

The age of the users had a significant influence on the system performance. By comparing the percentage of successful recognition attempts of Age 1 and Age 2, it is notable that the older participant obtained lower performance scores by interacting with D2 and D3. This can be caused by different factors. Firstly, the location of the biometric sensor: the lateral and the backside of the device made very uncomfortable the recognition process. This was confirmed by the elderly users along with the satisfaction survey. In fact, the user rated with a low score (2/5) the comfort, the time, and the easiness of D1 and D3. Secondly, even the quality of the fingerprint could have impacted the performance of the system.

Accessibility issues also influenced the recognition process. When people affected by cognitive and motor issues interacted with the system, they generally reached a lower level of performance compared with the Age 1 group. The accessibility of the scenario and the system were affected by the cognitive and motor issues of users. Just a few users belonging to the accessibility group were former costumer of mobile biometric solutions. Besides, due to their accessibility concerns, some participants could not interact with the system or complete the task required in the scenarios. As the elderly users, participants with accessibility issues had several interaction problems with the D2 and D3.

High percentages of incorrect interactions were reported when the user completes the second and third scenarios. The location of the fingerprint sensor made the user feel uncomfortable while using the D2 and D3. Thus, the D1 was rated as the favourite device by all the cognitive-issues groups.

Chapter 8 Conclusions and Recommendations for Future Works

This is the last chapter of the Thesis which discusses its main contributions. Across the next sections, the main outcomes reached thanks to our work will be listed and, besides, useful recommendations for future works will be provided.

8.1 Thesis outcome

This work established a formal methodology to conduct more traceable and comparable user interaction evaluations in biometrics. The starting point was state-of-the-art in the usability and user interaction evaluation. Hence, the main goal was to

analyse the previous methodologies and methods proposed to evaluate the usability and the influence of the user interaction on the biometric recognition outcome.

The initial step took us to the conclusion that very few studies fully considered the whole characteristics of the user and no methodology included the accessibility in biometric user interaction assessments. Whether or not a user has the possibility to interact with a specific biometric system is related to accessibility. Some characteristics of the scenario and the system could not allow the user to interact with the biometric sensor. At the same time, the accessibility concerns of the user (e.g. motor or cognitive issues) make the biometric system unapproachable. Skipping the information regarding the accessibility means not considering an important aspect related to the user interaction.

The accessibility evaluation helps to have full knowledge of all those aspects that influence the interaction between the user and the biometric systems. For this reason, providing an accessibility methodology enhances the reliability and the traceability of biometric user interaction assessments.

The main contribution of this thesis was to present a formal methodology to report the accessibility in biometric user interaction evaluations (described in Chapter 4). This methodology sets specific metrics to evaluate the accessibility of the scenario, of the system, and the impact of accessibility concerns on the recognition outcome.

According to this methodology, we reported the results obtained by analysing the data collected during two user interaction assessments (Chapters 6 and 7).

The results obtained carrying out our analysis allowed us to define the following considerations:

- Having access to a scenario or to a biometric system is strictly tied to the experience, and health status of the users.
- The lack of experience and the low dexterity of elderly users cause several interaction problems concerning the accessibility to the scenario and the biometric system.
- Accessibility issues and aging affect the outcome of the recognition system under different aspects: performance, usability, and sample quality.

- The behaviour of the user, while performing a biometric authentication, changes depending on his group sectors.
- The users' preference regarding the configuration of the biometric system depends on the subject group sectors.
- Biometric applications can be applied to improve the accessibility in several daily tasks required in private and public contexts.

8.2 Recommendations for Future Work

Concerning the results obtained, we can provide the following guidelines for future studies:

- When enrolling accessibility groups, it is a good practice to specify the motor or the cognitive problems affecting the users' capabilities and possibilities. Each accessibility issue influences the recognition process under different points of view.
- Justifying the reasons of no interactions (if it depends on the user, on the system, or the scenario) is recommendable to understand which specific factors prevent user form interacting with biometric recognition devices.
- Elderly users must be included while evaluating the accessibility of biometric systems. As demonstrated by our studies, there is a strong correlation between the user's age and the accessibility to the biometric process.
- Carrying out long term-evaluations (considering more than 2 sessions) is recommended to study the degree with which users improve their experience and interaction with biometric applications.
- A more careful study of users' behaviour while testing biometric applications. Besides the subjects' emotions and expressions, it could be helpful to establish if even the gesture of approaching biometric sensors depends on the user's characteristics (both age and health status). When users interact with mobile biometrics apps, it could be interesting to establish if people, belonging to the same user sector, approach the biometric sensor, or hold the smartphone in the same way.

- Implementing biometric solutions to support the user in daily tasks and testing them through accessibility evaluations is necessary to promote biometrics among more and more categories of users and, besides, to instruct users about a more conscious use of biometrics in everyday scenarios.

Reference

[1] "Facial recognition to feature on over 800m mobiles by 2024 | Planet Biometrics News." [Online]. Available: https://www.planetbiometrics.com/article-details/i/10722/.

[2] R. Blanco-Gonzalo, R. Sanchez-Reillo, O. Miguel-Hurtado, and J. Liu-Jimenez, "Usability analysis of dynamic signature verification in mobile environments," in *Biometrics Special Interest Group (BIOSIG), 2013 International Conference of the*, 2013, pp. 1–9.

[3] R. Blanco-Gonzalo, R. Sanchez-Reillo, R. Ros-Gomez, and B. Fernandez-Saavedra, "User acceptance of planar semiconductor fingerprint sensors," in *Proceedings - International Carnahan Conference on Security Technology*, 2016, vol. 2015-January, pp. 31–36.

[4] "ISO - ISO/IEC 19795-2:2007 - Information technology — Biometric performance testing and reporting — Part 2: Testing methodologies for technology and scenario evaluation." [Online]. Available: https://www.iso.org/standard/41448.html.

[5] "ISO/IEC AWI 21472 - Information technology -- Scenario evaluation methodology for user interaction influence in biometric system performance." [Online]. Available: https://www.iso.org/standard/70950.html.

[6] "ISO - International Organization for Standardization." [Online]. Available: https://www.iso.org/home.html.

[7] "National Institute of Standards and Technology | NIST." [Online]. Available: https://www.nist.gov/.

[8] E. P. Kukula, M. J. Sutton, and S. J. Elliott, "The humanbiometric-sensor interaction evaluation method: Biometric performance and usability measurements," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 4, pp. 1–8, Apr. 2010.

[9] "Fingerprint Recognition Through Circular Sampling." [Online]. Available: https://www.cis.rit.edu/research/thesis/bs/1999/chang/thesis.html..

[10] P. S. Prasad, B. Sunitha Devi, M. Janga Reddy, and V. K. Gunjan, "A survey of fingerprint recognition systems and their applications," in *Lecture Notes in Electrical Engineering*, 2019, vol. 500, pp. 513–520.

[11] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, "Face recognition systems: A survey," *Sensors (Switzerland)*, vol. 20, no. 2, 2020.

[12] J. Križaj, V. Štruc, and N. Pavešić, "Adaptation of SIFT features for robust face recognition," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2010, vol. 6111 LNCS, no. PART 1, pp. 394–404.

[13] K. Kim, "Face Recognition using Principle Component Analysis."

[14] "ISO - ISO/IEC JTC 1/SC 37 - Biometrics." [Online]. Available: https://www.iso.org/committee/313770.html.

[15] "Apple's new iPhone will read your fingerprint - The Verge." [Online]. Available: https://www.theverge.com/2013/9/10/4715372/confirmed-apple-iphone-5s-will-include-touch-id-fingerprint-scanner..

[16] "The future is here: iPhone X - Apple." [Online]. Available: https://www.apple.com/newsroom/2017/09/the-future-is-here-iphone-x/.

[17] "ISO - ISO/IEC 29794-1:2016 - Information technology — Biometric sample quality — Part 1: Framework." [Online]. Available: https://www.iso.org/standard/62782.html.

[18] "ISO - ISO/IEC 19795-1:2006 - Information technology — Biometric performance testing and reporting — Part 1: Principles and framework." [Online]. Available: https://www.iso.org/standard/41447.html.

[19] "ISO - ISO/IEC TR 19795-3:2007 - Information technology — Biometric performance testing and reporting — Part 3: Modality-specific testing." [Online]. Available: https://www.iso.org/standard/41449.html.

[20] "ISO - ISO/IEC 19795-4:2008 - Information technology — Biometric performance testing and reporting — Part 4: Interoperability performance testing." [Online]. Available: https://www.iso.org/standard/46329.html.

[21] "ISO - ISO/IEC 19795-5:2011 - Information technology — Biometric performance testing and reporting — Part 5: Access control scenario and grading scheme." [Online]. Available: https://www.iso.org/standard/51768.html.

[22] "ISO - ISO/IEC 19795-6:2012 - Information technology — Biometric performance testing and reporting — Part 6: Testing methodologies for operational evaluation." [Online]. Available: https://www.iso.org/standard/50873.html.

[23] "ISO - ISO/IEC 19795-7:2011 - Information technology — Biometric performance testing and reporting — Part 7: Testing of on-card biometric comparison algorithms." [Online]. Available: https://www.iso.org/standard/53059.html.

[24] "ISO - ISO/IEC TS 19795-9:2019 - Information technology — Biometric performance testing and reporting — Part 9: Testing on mobile devices." [Online]. Available: https://www.iso.org/standard/78101.html.

[25] International Standardization Organization (ISO), "ISO/IEC 19795-1:2006. Information technology -- Biometric performance testing and reporting -- Part 1: Principles and framework." .

[26] B. C. Stanton, M. F. Theofanos, S. M. Furman, J. M. Libert, S. Orandi, and J. D. Grantham, “Usability testing of a contactless fingerprint device: part 1,” Gaithersburg, MD, Dec. 2016.

[27] B. Stanton, M. Theofanos, S. Furman, P. J. Grother, P. Grother, and P. Pritzker, “Usability Testing of a Contactless Fingerprint Device: Part 2,” 2016.

[28] S. M. Furman, B. C. Stanton, M. F. Theofanos, J. M. Libert, and J. D. Grantham, “Contactless fingerprint devices usability test,” Gaithersburg, MD, Mar. 2017.

[29] N. Gunson, D. Marshall, F. McInnes, and M. Jack, “Usability evaluation of voiceprint authentication in automated telephone banking: Sentences versus digits,” *Interact. Comput.*, vol. 23, no. 1, pp. 57–69, Jan. 2011.

[30] R. Blanco-Gonzalo, O. Miguel-Hurtado, R. Sanchez-Reillo, and A. Gonzalez-Ramirez, “Usability analysis of a handwritten signature recognition system applied to mobile scenarios,” in *2013 47th International Carnahan Conference on Security Technology (ICCST)*, 2013, pp. 1–6.

[31] M. F. Theofanos, B. Stanton, C. Sheppard, and R. Micheals, “Usability Testing of Face Image Capture for US Ports of Entry,” in *2008 IEEE Second International Conference on Biometrics: Theory, Applications and Systems*, 2008, pp. 1–6.

[32] “ISO - ISO/IEC 29794-4:2017 - Information technology — Biometric sample quality — Part 4: Finger image data.” [Online]. Available: https://www.iso.org/standard/62791.html.

[33] “ISO/IEC TR 29794-5:2010 - Information technology -- Biometric sample quality -- Part 5: Face image data.” [Online]. Available: https://www.iso.org/standard/50912.html.

[34] “Biometric Quality Homepage | NIST.” [Online]. Available: https://www.nist.gov/programs-projects/biometric-quality-homepage.

[35] S. K. Modi and S. J. Elliott, “Impact of Image Quality on Performance: Comparison of Young and Elderly Fingerprints.”

[36] “Development of NFIQ 2.0 | NIST.” [Online]. Available: https://www.nist.gov/services-resources/software/development-nfiq-20.

[37] “ISO 9241-11:1998 - Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability.” [Online]. Available: https://www.iso.org/standard/16883.html.

[38] “Visualization and Usability Group | NIST.” [Online]. Available: https://www.nist.gov/itl/iad/visualization-and-usability-group.

[39] R. J. Micheals, B. Stanton, M. Theofanos, S. Orandi, C. M. Gutierrez, and W. Jeffrey, “A Taxonomy of Definitions for Usability Studies in Biometrics,” 2006.

[40] “Does Habituation Affect Fingerprint Quality?,” 2006.

[41] N. Visualization and U. Group, “Usability & Biometrics Ensuring Successful Biometric Systems,” 2008.

[42] "Homeland Security | Home." [Online]. Available: https://www.dhs.gov/..

[43] "Federal Register :: United States Visitor and Immigrant Status Indicator Technology Program ('US-VISIT'); Enrollment of Additional Aliens in US-VISIT; Authority To Collect Biometric Data From Additional Travelers and Expansion to the 50 Most Highly Trafficked Land Border Ports of Entry." [Online]. Available: https://www.federalregister.gov/documents/2008/12/19/E8-30095/united-states-visitor-and-immigrant-status-indicator-technology-program-us-visit-enrollment-of.

[44] B. Stanton, M. Theofanos, S. Orandi, R. Micheals, and N. F. Zhang, "Effects of Scanner Height on Fingerprint Capture," *NIST Rep.*, vol. 51, no. 10, pp. 592–596, 2007.

[45] B. C. Stanton, M. F. Theofanos, S. Orandi, R. J. Micheals, and N. F. Zhang, "Usability Testing of Ten-Print Fingerprint Capture." 08-Oct-2007.

[46] M. Theofanos *et al.*, "Usability Testing of Height and Angles of Ten-Print Fingerprint Capture NISTIR 7504 Usability Testing of Height and Angles of Ten-Print Fingerprint Capture," 2008.

[47] M. Theofanos *et al.*, "Assessing Face Overlay."

[48] R. Blanco-Gonzalo, L. Diaz-Fernandez, O. Miguel-Hurtado, and R. Sanchez-Reillo, "Usability Evaluation of Biometrics in Mobile Environments," Springer, Cham, 2014, pp. 289–300.

[49] M. Collotta, V. Conti, M. Collotta, G. Pau, and S. Vitabile, "Usability Analysis of a Novel Biometric Authentication Approach for Android-based Mobile Devices," *J. Telecommun. Inf. Technol.*, vol. 4, no. October 2015, 2014.

[50] M. Boakes, R. Guest, F. Deravi, and B. Corsetti, "Exploring Mobile Biometric Performance Through Identification of Core Factors and Relationships," *IEEE Trans. Biometrics, Behav. Identity Sci.*, vol. 1, no. 4, pp. 278–291, Sep. 2019.

[51] "IEEE Xplore Full-Text PDF:" [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5422748&casa_token=a9Kd_mB-CiUAAAAA:S3PPdAPjy-n2nzPmKK6TaMTk_jXC_5HzxCnrsEeFaXj6-LsBIHB13Qd30eLQQz_hX7pwsrPujA&tag=1.

[52] "IEEE Xplore Full-Text PDF:" [Online]. Available: https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5678710&casa_token=kOrch0HPSOEAAAAA:MbqSsN5d8bJRTt0XB8vVNEVMTBO_FleBPL89oYa82dVtG-zR4OFA0n-64YncPFWlMtA86fi_ZQ.

[53] O. Miguel-Hurtado, R. Blanco-Gonzalo, R. Guest, and C. Lunerti, "Interaction evaluation of a mobile voice authentication system," in *2016 IEEE International Carnahan Conference on Security Technology (ICCST)*, 2016, pp. 1–8.

[54] O. Miguel-Hurtado, R. Guest, and C. Lunerti, "Voice and face interaction evaluation of a mobile authentication platform," in *2017 International Carnahan*

Conference on Security Technology (ICCST), 2017, pp. 1–6.

[55] "ISO 26800:2011 - Ergonomics -- General approach, principles and concepts." [Online]. Available: https://www.iso.org/standard/42885.html.

[56] R. Sanchez-Reillo, R. Blanco-Gonzalo, J. Liu-Jimenez, M. Lopez, and E. Canto, "Universal access through biometrics in mobile scenarios," *Security Technology (ICCST), 2013 47th International Carnahan Conference on*. pp. 1–6, 2013.

[57] B. Stanton, M. Theofanos, and C. Sheppard, "A Study of Users with Visual Disabilities and a Fingerprint Process," 2008.

[58] R. Blanco-Gonzalo, R. Sanchez-Reillo, C. Sanchez-Redondo, and J. L. Alonso-Aguilera, "Accessibility evaluation of a mobile biometric recognition system," in *2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*, 2016, pp. 1–6.

[59] B. Corsetti, R. And Sanchez-Reillo, R. Guest, M. Santopietro, R. Sanchez-Reillo, and R. M. Guest, *Face Image Analysis in Mobile Biometric Accessibility Evaluations*. 2019.

[60] R. Blanco-Gonzalo, R. Sanchez-Reillo, L. Martínez-Normand, B. Fernandez-Saavedra, and J. Liu-Jimenez, "Accessible Mobile Biometrics for Elderly."

[61] C. Riley, H. McCracken, and K. Buckner, "Fingers, veins and the grey pound," 2007, p. 149.

[62] R. Blanco-Gonzalo, C. Lunerti, R. Sanchez-Reillo, and R. M. Guest, "Biometrics: Accessibility challenge or opportunity?," *PLoS One*, vol. 13, no. 3, p. e0194111, Mar. 2018.

[63] "General Data Protection Regulation (GDPR) – Official Legal Text." [Online]. Available: https://gdpr-info.eu/.

[64] "DigitalPersona EikonTouch 710 fingerprint reader." [Online]. Available: https://www.neurotechnology.com/fingerprint-scanner-digitalpersona-eikontouch-710.html.

[65] "AXIS M1011 Network Camera | Axis Communications." [Online]. Available: https://www.axis.com/products/axis-m1011.

[66] "OnePlus 3T - Technical Specification - OnePlus (United Kingdom)." [Online]. Available: https://www.oneplus.com/uk/support/spec/oneplus-3t.

[67] "OpenCV." [Online]. Available: https://opencv.org/.

[68] "Xperia XZ Premium | Android smart phone by Sony | Sony UK." [Online]. Available: https://www.sony.co.uk/electronics/smartphones/xperia-xz-premium.

[69] "Welcome to Neffos." [Online]. Available: https://www.neffos.in/.

[70] "Freeware SDK and .NET components for biometric application development." [Online]. Available: https://www.neurotechnology.com/free-fingerprint-verification-sdk.html.

[71] “Fingerprint, face, eye iris, voice and palm print identification, speaker and object recognition software.” [Online]. Available: https://www.neurotechnology.com/.

[72] “VeriLook face identification technology, algorithm and SDK for PC, smartphones and Web.” [Online]. Available: https://www.neurotechnology.com/verilook.html.

[73] “One-way ANOVA - An introduction to when you should run this test and the test hypothesis | Laerd Statistics.” [Online]. Available: https://statistics.laerd.com/statistical-guides/one-way-anova-statistical-guide.php.

[74] “Fingerprint Minutiae Viewer (FpMV) | NIST.” [Online]. Available: https://www.nist.gov/services-resources/software/fingerprint-minutiae-viewer-fpmv.

ANNEX 1: Consent form for data storage

University Group of Identification Technologies (GUTI) **Department of Electronic Technology**
CONSENT TO CAPTURE DATA FOR EVALUATIONS CARRIED OUT BY THE UNIVERSITY GROUP OF IDENTIFICATION TECHNOLOGIES (GUTI) OF THE CARLOS III UNIVERSITY OF MADRID

In the University Group of Identification Technologies (GUTI) of the Carlos III University of Madrid, different biometric evaluations are carried out. During the evaluations, volunteers are requested to participate in. This document aims to ask your permission to add your contacts and identification data in our database. With this consent, you can decide if the GUTI can use this data to inform you about future evaluations. Otherwise, your contacts will be deleted once the evaluation will be finished.

The following table details the information related to this consent:

Responsible	Universidad Carlos III de Madrid	https://www.uc3m.es/ss/Satellite/UC3MInstitucional/es/TextoDosColumnas/1371250765889/Proteccion_de_Datos
Legitimation	Legitimation Consent of the interested party	https://www.boe.es/doue/2016/119/L00001-00088.pdf
Purpose	Record your biographical data to facilitate the correct labeling of captured biometric data.	Optionally, if you wish, they will be used to inform you of future evaluations made in the GUTI, in case you are interested in participating.
Data to be requested	- Name and Surname - DNI / NIE / Passport Number - Country of birth - Country of Residence - Birthdate - Email address - Telephone contact - Laboral sector	
Data transfer	No data will be transferred to third parties	
Rights	**Access or rectification**: to consult and/or request the modification of your data. **Deletion**: to request the deletion of your data. **Opposition or limitation**: to request that they not be treated or that a limitation be established in their treatment. **Portability**: to request the transmission of your data to a third party.	The interested party must send an email to protdatos@uc3m.es with his/her name, surnames, and DNI/NIE indicating what right he/she wishes to exercise on the database **CONTACTOS_SUJETOS_GUTI**

Consent:

I_____, with ID Card Number: _____, I declare to have read and understood the conditions that are detailed in this document **I authorize** to include my data in the database, for the correct labeling of the biometric samples captured by the GUTI.

	YES	NO
Additionally, I authorize my data to remain in the possession of the GUTI to inform me of future evaluations that the group		

The absence of marks in any of the boxes will imply a NO. Likewise, the presence of marks in both boxes will imply a NO.

Date: _____

Signature: _____

ANNEX 2: Information Document about the experiment

University Group of Identification Technologies (GUTI) **Department of Electronic Technology**
CONSENT TO COLLECT BIOMETRIC DATA WITH REFERENCE: GUTI_ControlAccesso_1_2018

You are being invited to take part in a research project on biometric identification. The aim is to analyse the performance and accessibility of an access control system through biometric recognition using fingerprint and face samples.

During this evaluation will be collected all information from the biometric samples (face and fingerprint) and the users' opinion (through questionnaires).

Before deciding to participate in this project, it is important that you understand the reasons for our work and the data collection process. Please, take your time to read all the information that is detailed in the following table. Do not hesitate to ask the person who will guide you during the process if you find something that is not clear enough or if you need more information.

The entire process will not involve any risk to the user.

Responsible	Universidad Carlos III de Madrid	https://www.uc3m.es/ss/Satellite/UC3MInstitucional/es/TextoDosColumnas/1371250765889/Proteccion_de_Datos
Legitimation	Consent of the interested party.	https://www.boe.es/doue/2016/119/L00001-00088.pdf
Purpose	Evaluate the performance and accessibility of an access control system through biometric recognition by fingerprint and face.	The phases of this research project will be detailed at the end of this document.
Personal and biometric data that will be requested	- Prior knowledge of technology - Previous knowledge of biometric recognition - Images of face - Fingerprint images	These personal and biometric data will be kept in the database for a minimum period of one year. After this period time, if you have not exercised your right to deletion of the data, the entire database will be deleted.
Data transfer	No data will be shared with third parties.	
Rights	**Access or rectification**: to consult and/or request the modification of your data. **Deletion**: to request the deletion of your data.	Every participant can send an email to protdatos@uc3m.es with his name, surnames, and DNI/NIE indicating what right he wants to exercise on the database **GUTI_ControlAccesso_1_2018**

	Opposition or limitation: to request that they not be treated or that a limitation be established in their treatment. **Portability**: to request the transmission of your data to a third party.	

The process of biometric samples' capturing is divided into two phases separated by one week time:

- Visit 1

- Explanation of the purpose of the project
- Delivery of consent forms following the GDPR.
- Questionnaire on the use of biometric systems
- Collection of contact data and biometric samples

- Visit 2

- Collection of biometric samples.
- Questionnaire about your opinion regarding the devices used.

Consent:

I_____, with ID Card Number: _____, I declare to have read and understood the conditions that are detailed in this document and **I authorize** the University Group of Identification Technologies (GUTI) to collect my personal and biometric data (face images and fingerprint images) during the research project described above.

Date: _____

Signature: _____